Beyond the Dirac phase factor: Dynamical Quantum Phase-Nonlocalities in the Schrödinger Picture

Konstantinos Moulopoulos*

University of Cyprus, Department of Physics, 1678 Nicosia, Cyprus

(May 10, 2011)

Abstract

Generalized solutions of the standard gauge transformation equations are presented and discussed in physical terms. They go beyond the usual Dirac phase factors and they exhibit nonlocal quantal behavior, with the well-known Relativistic Causality of classical fields affecting directly the *phases* of wavefunctions in the Schrödinger Picture. These nonlocal phase behaviors, apparently overlooked in path-integral approaches, give a natural account of the dynamical nonlocality character of the various (even static) Aharonov-Bohm phenomena, while at the same time they seem to respect Causality. Indeed, for particles passing through nonvanishing magnetic or electric fields they lead to cancellations of Aharonov-Bohm phases at the observation point, generalizing earlier semiclassical experimental observations (of Werner & Brill) to delocalized (spread-out) quantum states. This leads to a correction of previously unnoticed sign-errors in the literature, and to a natural explanation of the deeper reason why certain time-dependent semiclassical arguments are consistent with static results in purely quantal Aharonov-Bohm configurations. These nonlocalities also provide a remedy for misleading results propagating in the literature (concerning an uncritical use of Dirac phase factors, that persists since the time of Feynman's work on path integrals). They are shown to conspire in such a way as to exactly cancel the instantaneous Aharonov-Bohm phase and recover Relativistic Causality in earlier "paradoxes" (such as the van Kampen thought-experiment), and to also complete Peshkin's discussion of the electric Aharonov-Bohm effect in a *causal* manner. The present formulation offers a direct way to address time-dependent single- *vs* double-slit experiments and the associated causal issues – issues that have recently attracted attention, with respect to the inability of current theories to address them.

*Electronic address: cos@ucy.ac.cy

I. INTRODUCTION

The Dirac phase factor − with a phase containing spatial or temporal integrals of potentials (of the general form $\int^{r} \boldsymbol{A} \cdot d\boldsymbol{x}' - c\int^{t} \phi dt'$) − is the standard and widely used solution of the gauge transformation equations of Electrodynamics (with $\boldsymbol{A}$ and ϕ vector and scalar potentials respectively). In a quantum mechanical context, it connects wavefunctions of two systems (with different potentials) that experience the same classical fields at the observation point $(\boldsymbol{r}, t)$, the two more frequently discussed cases being: either systems that are completely gauge-equivalent (a trivial case with no physical consequences), or systems that exhibit phenomena of the Aharonov-Bohm type (magnetic or electric)[1] − and then this Dirac phase has nontrivial observable consequences (mathematically, this being due to the fact that the corresponding "gauge function" is now multiple-valued). In the above two cases, the classical fields experienced by the two (mapped) systems are equal *at every point* of the accessible spacetime region. However, it has not been widely realized that the gauge transformation equations, viewed in a more general context, can have *more general solutions* than simple Dirac phases, and these lead to wavefunction-*phase-nonlocalities* that have been widely overlooked and that seem to have important physical consequences. These nonlocal solutions are applicable to cases where the two systems are allowed to experience *different fields* at spacetime points (or regions) that are *remote* to (and do *not* contain) the observation point $(\boldsymbol{r}, t)$. In this article we rigorously show the existence of these generalized solutions, demonstrate them in simple physical examples, and fully explore their structure, presenting cases (and closed analytical results for the wavefunction-phases) that actually connect (or map) two quantal systems that are **neither physically equivalent nor of the usual Aharonov-Bohm type**. We also fully investigate the consequences of these generalized (*nonlocal*) influences (on wavefunction-phases) and find them to be numerous and important; we actually find them to be of a different type in static and in time-dependent field-configurations (and in the latter cases we show that they lead to Relativistically *causal* behaviors, that apparently resolve earlier "paradoxes" arising in the literature from the use of standard Dirac phase factors). The nonlocal phase behaviors discussed in the present work may be viewed as a justification for the (recently emphasized[2]) terminology of "dynamical nonlocalities" associated with all Aharonov-Bohm effects (even static ones), although in our

approach these nonlocalities seem to also respect Causality (without the need to independently invoke the Uncertainty Principle) – and, to the best of our knowledge, this is the first theoretical picture with such characteristics.

In order to introduce some background and further motivation for this article let us first remind the reader of a very basic property that will be central to everything that follows, which however is usually taken to be valid only in a restricted context (but is actually more general than often realized). This property is a simple (U(1)) phase-mapping between quantum systems, and is usually taken in the context of gauge transformations, ordinary or singular**;** here, however, it will appear in a more general framework, hence the importance of reminding of its independent, basic and more general origin. We begin by recalling that, if $\Psi_1(\mathbf{r},t)$ and $\Psi_2(\mathbf{r},t)$ are solutions of the time-dependent Schrödinger (or Dirac) equation for a quantum particle of charge q that moves (as a test particle) in two distinct sets of (predetermined and classical) vector and scalar potentials $(\mathbf{A}_1,\phi_1)$ and $(\mathbf{A}_2,\phi_2)$, that are generally spatially- and temporally-dependent [and such that, at the spacetime point of observation $(\mathbf{r},t)$, the magnetic and electric fields are the same in the two systems], then we have the following formal connection between the solutions (wavefunctions) of the two systems

$$\Psi_2(\mathbf{r},t) = e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}\Psi_1(\mathbf{r},t), \tag{1}$$

with the function $\Lambda(\mathbf{r},t)$ required to satisfy

$$\nabla\Lambda(\mathbf{r},\mathbf{t}) = \mathbf{A}_2(\mathbf{r},t) - \mathbf{A}_1(\mathbf{r},t) \qquad \textit{and} \qquad -\frac{1}{c}\frac{\partial\Lambda(\mathbf{r},t)}{\partial t} = \phi_2\,(\mathbf{r},t) - \phi_1(\mathbf{r},t). \tag{2}$$

The above property can be immediately proven by substituting each Ψ_i into its corresponding (i_{th}) time-dependent Schrödinger equation (namely with the set of potentials $(\mathbf{A}_i(\mathbf{r},t),\phi_i(\mathbf{r},t))$)**:** one can then easily see that (1) and (2) guarantee that both Schrödinger equations are indeed satisfied together (after cancellation of a global phase factor in system 2, see Appendix A for a detailed proof). [In addition, the equality of all classical fields at the observation point, namely $\mathbf{B}_2(\mathbf{r},t) = \nabla\times\mathbf{A}_2(\mathbf{r},t) = \nabla\times\mathbf{A}_1(\mathbf{r},t) = \mathbf{B}_1(\mathbf{r},t)$ for the magnetic fields and $\mathbf{E}_2(\mathbf{r},t) = -\nabla\phi_2\,(\mathbf{r},t) - \frac{1}{c}\frac{\partial\mathbf{A}_2(\mathbf{r},t)}{\partial t} = -\nabla\phi_1\,(\mathbf{r},t) - \frac{1}{c}\frac{\partial\mathbf{A}_1(\mathbf{r},t)}{\partial t} = \mathbf{E}_1(\mathbf{r},t)$ for the electric fields, is obviously consistent with all equations (2) (as is easy to see if we take the *curl* of the 1st and the *grad* of the 2nd) – provided, at least, that $\Lambda(\mathbf{r},t)$ is such that

interchanges of partial derivatives with respect to all spatial and temporal variables (at the point $(\mathbf{r}, t)$) are allowed].

As already mentioned, the above fact is of course well-known within the framework of the theory of quantum mechanical gauge transformations (the usual case being for $\mathbf{A}_1 = \phi_1 = 0$, hence for a mapping from a system with no potentials); but in that framework, these transformations are supposed to connect (or map) two *physically equivalent systems* (more rigorously, this being true for ordinary gauge transformations, in which case the function $\Lambda(\mathbf{r}, t)$, the so-called gauge function, is unique (single-valued) in spacetime coordinates). In a formally similar manner, the above argument is also often used in the context of the so-called "singular gauge transformations", where Λ is multiple-valued, but the above equality of classical fields is still imposed (at the observation point, which always lies in a physically accessible region); then the above simple phase mapping (at all points of the physically accessible spacetime region, that always and everywhere experience equal fields) leads to the standard phenomena of the Aharonov-Bohm type, reviewed later below, where *unequal fields in physically-inaccessible regions* have observable consequences. However, we should keep in mind that that above property ((1) and (2) taken together) can be *more generally valid* $-$ and in this article we will present cases (and closed analytical results for the appropriate phase connection $\Lambda(\mathbf{r}, t)$) that actually connect (or map) two systems (in the sense of (1)) that are *neither physically equivalent nor of the usual Aharonov-Bohm type.* And naturally, because of the above provision of field equalities at the observation point, it will turn out that any nonequivalence of the two systems will involve *remote* (although *physically accessible*) regions of spacetime, namely regions that do *not* contain the observation point $(\mathbf{r}, t)$ (and in which regions, as we shall see, the classical fields experienced by the particle may be *different* in the two systems).

II. MOTIVATION

One may wonder on the actual reasons why one should be looking for more general cases of a simple phase mapping of the type (1) between *nonequivalent* systems. To answer this, let us take a step back and first recall some simple and well-known results that originate from the above phase mapping. It is standard knowledge, for example, that, if we want to find solutions $\Psi(x, t)$ of the t-dependent Schrödinger (or Dirac) equation for a quantum

particle (of charge q) that moves along a (generally curved) one-dimensional (1-D) path, and in the presence (somewhere in the embedding 3-dimensional (3-D) space) of a fairly localized (and time-independent) classical magnetic flux Φ that *does **not** pass through any point of the path,* then we formally have

$$\Psi(x,t)^{(\Phi)} \sim e^{i\frac{q}{\hbar c}\int_{x_0}^{x}\mathbf{A}(\mathbf{r}')\cdot d\mathbf{r}'}\Psi(x,t)^{(0)} \tag{3}$$

(the dummy variable $\mathbf{r}'$ describing points along the 1-D path, and the term "formally" meaning that the above is valid *before imposition of any boundary conditions* (that are meant to be imposed only on the system with the flux Φ[3])). In (3), $\Psi(x,t)^{(0)}$ is a formal solution of the same system in the case of absence of any potentials (hence $\Phi = 0$ everywhere in the 3-D space). The above holds because, for *all points* $\mathbf{r}'$ of the 1-D path, the particle experiences a vector potential $\mathbf{A}(\mathbf{r}\prime)$ of the form $\mathbf{A}(\mathbf{r}\prime) = \nabla'\mathbf{\Lambda}(\mathbf{r}')$ (since the magnetic field is $\nabla'\times\mathbf{A}(\mathbf{r}') = 0$ for *all* $\mathbf{r}'$, by assumption), in combination with the above phase-mapping (with a phase $\frac{q}{\hbar c}\mathbf{\Lambda}(\mathbf{r})$) between two quantum systems, one in the presence and one in the absence of a vector potential (i.e. the potentials in (2) being $\mathbf{A}_1 = 0$ and $\mathbf{A}_2 = \mathbf{A}$, together with $\phi_2 = \phi_1 = 0$ if we decide to attribute everything to vector potentials only). In this particular system, the obvious $\mathbf{\Lambda}(\mathbf{r})$ that solves the above $\mathbf{A}(\mathbf{r}) = \nabla\mathbf{\Lambda}(\mathbf{r})$ (for **all** points of the 1-D space available to the particle) is indeed $\mathbf{\Lambda}(\mathbf{r}) = \mathbf{\Lambda}(\mathbf{r}_0) + \int_{\mathbf{r}_0}^{\mathbf{r}}\mathbf{A}(\mathbf{r}')\cdot d\mathbf{r}'$, and this gives (3) (if $\mathbf{r}$ denotes the above point x of observation and $\mathbf{r}_0$ the arbitrary initial point x_0 (both lying on the physical path), and if the constant $\mathbf{\Lambda}(\mathbf{r}_0)$ is taken to be zero).

What if, however, some parts of the magnetic field that comprise the magnetic flux Φ actually *pass through* some points or a whole region (interval) of the path available to the particle? In such a case, the above standard argument is not valid (as $\mathbf{A}$ cannot be written as a grad at any point of the interval where the magnetic field $\nabla\times\mathbf{A}\neq 0$). Are there however general results that we can still write for $\Psi^{(\Phi)}(x,t)$, if the spatial point of observation x is again outside the interval with the nonvanishing magnetic field? Or, what if in the previous problems, the magnetic flux (either remote, or partly passing through the path) is time-dependent $\Phi(t)$? (In that case then, there exists in general an additional electric field E induced by Faraday's law of Induction on points of the path, and the usual gauge transformation argument is once again not valid).

Returning to another standard (solvable) case (which is actually the "dual" or the "electric analog" of the above), if along the 1-D physical path the particle experiences only a

spatially-uniform (but generally time-dependent) classical scalar potential $\phi(t)$, we can again formally map $\Psi(x,t)^{(\phi)}$ to a potential-free solution $\Psi(x,t)^{(0)}$, through a $\Lambda(t)$ that now solves $-\frac{1}{c}\frac{\partial\Lambda(t)}{\partial t}=\phi(t)$, and this gives $\Lambda(t)=\Lambda(t_0)-c\int_{t_0}^{t}\phi(t')dt'$, leading to the "electric analog" of (3), namely

$$\Psi(x,t)^{(\phi)}\sim e^{-i\frac{q}{\hbar}\int_{t_0}^{t}\phi(t')dt'}\Psi(x,t)^{(0)} \tag{4}$$

with obvious notation. (Notice that, for either of the two mapped systems in this problem, the electric field is zero at all points of the path). What if, however, the scalar potential has also some x-dependence along the path (that leads to an electric field (in a certain interval) that the particle passes through)? In such a case, the above standard argument is again not valid. Are there however general results that we can still write for $\Psi(x,t)^{(\phi)}$, if the spatial point of observation x is again outside the interval with the nonvanishing electric field?

We state here directly that this article will provide affirmative answers to questions of the type posed above, by actually giving the corresponding general results in closed analytical forms.

At this point it is also useful to briefly reconsider the earlier mentioned case, namely of a time-dependent $\Phi(t)$ that is remote to the 1-D physical path, because in this manner we can immediately provide another motivation for the present work**:** this time-dependent problem is surrounded with a number of important misconceptions in the literature (the same being true about its electric analog, as we shall see)**:** the formal solution that is usually written down for a $\Phi(t)$ is again (3), namely the above spatial line integral of $\mathbf{A}$, in spite of the fact that $\mathbf{A}$ is now t-dependent**;** the problem then is that, because of the first of (2), Λ must now have a t-dependence and, from the second of (2), there must necessarily be scalar potentials involved in the problem (which have been by force set to zero, in our pre-determined mapping between vector potentials only). Having decided to use systems that experience only vector (and not scalar) potentials, the correct solution cannot be simply a trivial t-dependent extension of (3). A corresponding error is usually made in the electric dual of the above, namely in cases that involve $\mathbf{r}$-dependent scalar potentials, where (4) is still erroneously used, giving an $\mathbf{r}$-dependent Λ, although this would necessarily lead to the involvement of vector potentials (through the first of (2) and the $\mathbf{r}$-dependence of Λ) that have been neglected from the beginning $-$ a situation (and an error) that appears, in exactly this form, in the description of the so-called electric Aharonov-Bohm effect, as we shall see.

Speaking of errors in the literature, it might here be the perfect place to also point to the reader the most common misleading statement often made in the literature (and again, for notational simplicity, we restrict our attention to a one-dimensional system, with spatial variable x, although the statement is obviously generalizable to (and often made for systems of) higher dimensionality by properly using line integrals over arbitrary curves in space)**:** It is usually stated [e.g. in Brown & Holland[4], see i.e. their eq. (57) applied for vanishing boost velocity $\mathbf{v} = 0$] that the general gauge function that connects (through a phase factor $e^{i\frac{q}{\hbar c}\Lambda(x,t)}$) the wavefunctions of a quantum system with no potentials (i.e. with a set of potentials $(\mathbf{0}, 0)$) to the wavefunctions of a quantum system that moves in vector potential $\boldsymbol{A}(x,t)$ and scalar potential $\phi(x,t)$ (i.e. in a set of potentials $(\boldsymbol{A}, \phi)$) is the obvious combination (and a natural extension) of (3) and (4), namely

$$\Lambda(x,t) = \Lambda(x_0,t_0) + \int_{x_0}^{x} \boldsymbol{A}(x',t)dx' - c\int_{t_0}^{t} \phi(x,t')dt', \tag{5}$$

which, however, is **incorrect** for x and t uncorrelated variables**:** it does **not** satisfy the standard system of gauge transformation equations

$$\nabla\Lambda(x,t) = \mathbf{A}(x,t) \qquad \textbf{and} \qquad -\frac{1}{c}\frac{\partial\Lambda(x,t)}{\partial t} = \phi(x,t). \tag{6}$$

The reader can easily see why: (i) when the ∇ operator acts on eq.(5), it gives the correct $A(x,t)$ from the 1st term, but it also gives some annoying additional nonzero quantity from the 2nd term (that survives because of the x-dependence of ϕ); hence it invalidates the first of the basic system (6). (ii) Similarly, when the $-\frac{1}{c}\frac{\partial}{\partial t}$ operator acts on eq.(5), it gives the correct $\phi(x,t)$ from the 2nd term, but it also gives some annoying additional nonzero quantity from the 1st term (that survives because of the t-dependence of $\mathbf{A}$); hence it invalidates the second of the basic system (6). It is only when $\mathbf{A}$ is t-independent, and ϕ is spatially-independent, that eq.(5) is correct (as the above annoying terms do not appear and the basic system is satisfied). [An alternative form that is also given in the literature is again eq.(5), but with the variables that are not integrated over implicitly assumed to belong to the initial point (hence a t_0 replaces t in $\boldsymbol{A}$, and simultaneously an x_0 replaces x in ϕ). However, one can see again that the system (6) is not satisfied (the above differential operators, when acted on Λ, give $\boldsymbol{A}(x,t_0)$ and $\phi(x_0,t)$, hence not the values of the potentials at the point of observation (x,t) as they should), this not being an acceptable solution

either].

What is the problem here? Or, better put, what is the deeper reason for the above inconsistency? The short answer is the uncritical use of Dirac phase factors that come from path-integral treatments. It is indeed obvious that the form (5) that is often used in the literature (in canonical (non-path-integral) formulations where x and t are **uncorrelated** variables (and not correlated to produce a path $x(t)$)) *is not generally correct*, and that is one of the main points that has motivated this work. We will find *generalized results* that actually *correct* eq.(5) through extra nonlocal terms, and through the proper appearance of x_0 and t_0 (as in eq.(22) and eq.(23) to be found later in Section VII), and these are the **exact** ones (namely the exact $\Lambda(x,t)$, that at the end, upon action of ∇ and $-\frac{1}{c}\frac{\partial}{\partial t}$ satisfies exactly the basic system (6)). And the formulation that gives these results is generalized later in the article, for $\Lambda(x,y)$ (in the 2-D static case) and also for $\Lambda(x,y,t)$ (in the full dynamical 2-D case), and leads to the exact (nontrivial) forms of the phase function Λ that satisfy (in all cases) the system (6) $-$ with *the direct verification (i.e. proof, by "going backwards", that these forms are indeed the exact solutions) being given in the main text,* and with the rigorous mathematical derivations being presented in Appendices.

This article gives a full exploration of issues related to the above motivating discussion, by pointing to a "practical" (and generalized) use of gauge transformation mapping techniques, that at the end lead to these generalized (and, at first sight, unexpected) solutions for the general form of Λ. For cases such as the ones discussed above, or even more involved ones, there still appears to exist a simple phase mapping (between two inequivalent systems), but the phase connection Λ seems to contain not only integrals of potentials, but also "fluxes" of the classical fields from *remote* spacetime regions. The above mentioned systems are the simplest ones where these new results can be applied, but apart from this, the present investigation seems to lead to a number of nontrivial corrections of misleading (or even incorrect) reports of the above type in the literature, that are not at all marginal (and are due to an incorrect use of a path-integral viewpoint in an otherwise canonical framework). The generalized Λ-forms also lead to an honest resolution of earlier "paradoxes" (involving Relativistic Causality), and in some cases to a new interpretation of known semiclassical experimental observations, corrections of certain sign-errors in the literature, and nontrivial extensions of earlier semiclassical results to general (even completely delocalized) states. Most importantly, however, the new formulation seems capable of treating issues of Causality

in time-dependent slit experiments as we shall see[5].

III. THE STANDARD BACKGROUND

Having clearly stated that the above phase-mapping can be a more general property than is usually realized, let us briefly recall the history of the standard framework that such mappings have appeared. It was essentially from Weyl's work (1929), but also from independent proposals by Schrödinger (1922), Fock (1927) and London (1927)[6], that it was firmly established that a simple unitary (U(1)) phase factor certainly connects two quantum systems, when these are gauge-equivalent (and then the phase that connects their wavefunctions is basically the gauge function of an ordinary gauge transformation). A simple unitary phase-connection of this type is also reserved for quantum systems moving in multiple-connected spacetimes (with enclosed appropriately defined "fluxes" in the physically inaccessible regions) the corresponding "gauge transformation" termed singular, and the corresponding "gauge function" now being multiple-valued (although the wavefunctions of the "final" (mapped) system are still single-valued) leading to phenomena of the Aharonov-Bohm type[7]. As we already stressed, in the present work we report on a phase-connection between systems that are *not* "equivalent" (in the sense of the above two), since they can go through *different* classical fields in remote regions of space and/or time, and we give explicit forms of the appropriate "gauge functions". The results are exact, in analytical form, and they generalize the standard Dirac phase factors derived from path integral treatments (that are very often used in an incorrect way as we shall demonstrate); apart from a discussion of such misconceptions propagating in the literature, we also give actual applications of the new results in static and time-dependent experiments involving quantum charged particles inside electromagnetic potentials, both of the Aharonov-Bohm type (i.e. with inaccessible fields and their fluxes) but also with the particles actually passing through classical magnetic and electric fields, and even being in completely general (spead-out) quantum states (and not necessarily narrow wavepackets in semiclassical motion as typically done in the literature).

In Section I we recalled the mapping (1) and (2), together with an outline of its proof. Let us here briefly restate it for completeness and for a better flow of the arguments that will follow: if $\Psi_1(\mathbf{r},t)$ and $\Psi_2(\mathbf{r},t)$ are solutions of the time-dependent Schrödinger (or Dirac)

equation for a quantum particle of charge q that moves (as a test particle) in two distinct sets of (predetermined and classical) vector and scalar potentials $(\mathbf{A}_1, \phi_1)$ and $(\mathbf{A}_2, \phi_2)$ [and such that, at the spacetime point of observation $(\mathbf{r}, t)$, the magnetic and electric fields are the same in the two systems], then we have the following formal connection between the solutions (wavefunctions) of the two systems

$$\Psi_2(\mathbf{r}, t) = e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}\Psi_1(\mathbf{r}, t), \tag{7}$$

(by "formal" connection meaning again that this is valid *before imposition of any boundary conditions*, and these will have to be applied only on *our* system, i.e. system 2 (not necessarily on system 1, see [3])). In (7) $\Lambda(\mathbf{r}, t)$ is a general function that must satisfy (2), which we here want to treat as a system of Partial Differential Equations (PDEs)

$$\nabla\Lambda(\mathbf{r}, \mathbf{t}) = \mathbf{A}_2(\mathbf{r}, t) - \mathbf{A}_1(\mathbf{r}, t) \qquad \textit{and} \qquad -\frac{1}{c}\frac{\partial\Lambda(\mathbf{r}, t)}{\partial t} = \phi_2\,(\mathbf{r}, t) - \phi_1(\mathbf{r}, t). \tag{8}$$

In the static case, and if, for simplicity, we start from system 1 being completely free of potentials ($\mathbf{A}_1 = \phi_1 = 0$), the wavefunctions of the particle in system 2 (moving only in a static vector potential $A(\mathbf{r})$) will acquire an extra phase with an appropriate "gauge function" $\Lambda(\mathbf{r})$ that must satisfy

$$\nabla\Lambda(\mathbf{r}) = \mathbf{A}(\mathbf{r}). \tag{9}$$

As mentioned in Section II, the standard (and widely-used) solution of this is the line integral

$$\Lambda(\mathbf{r}) = \Lambda(\mathbf{r_0}) + \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{A}(\mathbf{r}').d\mathbf{r}' \tag{10}$$

(which, by considering two paths encircling an enclosed inaccessible magnetic flux, formally leads to the well-known magnetic Aharonov-Bohm effect[1]). It should however be stressed again that the above is only true if (9) is valid for **all** points $\mathbf{r}$ of the region where the particle moves, i.e. if the particle in system 2 moves (as a narrow wavepacket) always outside magnetic fields ($\nabla \times \mathbf{A} = 0$ everywhere) as already emphasized in Section II. Similarly, if the particle in system 2 moves only in a spatially homogeneous scalar potential $\phi(t)$, the

appropriate Λ must satisfy

$$-\frac{1}{c}\frac{\partial\Lambda(t)}{\partial t}=\phi(t), \tag{11}$$

the standard solution being

$$\Lambda(t)=\Lambda(t_0)-c\int_{t_0}^{t}\phi(t')dt' \tag{12}$$

that gives the extra phase acquired by system 2 (this result formally leading to the electric Aharonov-Bohm effect[1, 8] by applying it to two equipotential regions, such as two metallic cages held in distinct time-dependent scalar potentials). Once again, it should be stressed that the above is only true if (11) (and the assumed spatial homogeneity of the scalar potential ϕ and of Λ) is valid at **all** times t of interest, i.e. if the particle in system 2 moves (as a narrow wavepacket) always outside electric fields ($\mathbf{E}=-\nabla\phi-\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}=0$ at all times). (In the electric Aharonov-Bohm setup, the above is ensured by the fact that t lies in an interval of a finite duration T for which the potentials are turned on, in combination with the narrowness of the wavepacket**;** this guarantees that, during T, the particle has vanishing probability of being at the edges of the cage where the potential starts having a spatial dependence. The reader is referred to Appendix B of Peshkin[8] that demonstrates the intricasies of the electric Aharonov-Bohm effect, to which we return with an important comment at the end of Section XII).

In the present work, we relax the above assumptions and present more general solutions of the system of PDEs (8), covering cases where the particle is *not* necessarily a narrow wavepacket (it can actually be in completely delocalized states) and is *not* excluded from *remote* regions (in spacetime) of nonvanishing (or, more generally, of unequal) fields (magnetic or electric), regions therefore that are actually accessible to the particle (hence non-Aharonov-Bohm cases $-$ or even combinations of spatial multiple-connectivity of the magnetic Aharonov-Bohm type, but simultaneous simple-connectivity in spacetime (i.e. in the $(x,t)-$plane)). We find analytically nonlocal influences of these remote fields on $\Lambda(\mathbf{r},t)$ (with $(\mathbf{r},t)$ always denoting the observation point in spacetime), and therefore on the phases of wavefunctions at $(\mathbf{r},t)$, that seem to have a number of important consequences**:** they provide (i) a natural justification of earlier or more recent experimental observations for semiclassical behavior in simple-connected space (when the particles pass through nonvanishing (or, more generally, unequal) magnetic or electric fields), and also extensions to more general cases of

delocalized (spread-out) quantum states, (ii) a nontrivial correction to misleading or even incorrect results that appear often in the literature (errors that are of a different type for static and for time-dependent situations), and (iii) a natural remedy for Causality "paradoxes" in time-dependent Aharonov-Bohm configurations. These dynamical nonlocalities of quantum mechanical phases seem to have escaped from all path-integral treatments. An extension of the method applied to Maxwell's equations governing the behavior of the fields (rather than to the equations that give the "gauge function" Λ) indicates that these *phase nonlocalities* demonstrate in part a causal propagation of phases of quantum mechanical wavefunctions in the Schrödinger Picture (and these can address causal issues in time-dependent single- *vs* double-slit experiments, an area that seems to have recently attracted attention[2, 9, 10]).

IV. EXAMPLE OF GENERALIZED SOLUTIONS IN STATIC CASES

By way of an example we immediately provide a simple result that will be found later (in Section X) for a static (x, y)-case (and for simple-connected space) that generalizes the standard Dirac phase (10), namely

$$\Lambda(x,y) = \Lambda(x_0,y_0) + \int_{x_0}^{x} A_x(x',y)dx' + \int_{y_0}^{y} A_y(\mathbf{x_0},y')dy' + \left\{ \int\limits_{y_0}^{y} dy' \int\limits_{x_0}^{x} dx' B_z(x',y') + g(x) \right\} \tag{13}$$

$$\textit{with } g(x) \textit{ chosen so that } \left\{ \int\limits_{y_0}^{y} dy' \int\limits_{x_0}^{x} dx' B_z(x',y') + g(x) \right\} \textbf{:} \text{ is independent of } x.$$

In the above $B_z = (\boldsymbol{B}_2 - \boldsymbol{B}_1)_z$ is the difference of perpendicular magnetic fields in the two systems, which can be nonvanishing at regions remote to the observation point (x, y) (see below). (It is reminded that at the point of observation $B(x, y) = 0$, signifying the essential fact that the fields in the two systems are identical (recall that $B_z = B_{2z} - B_{1z}$) at the point of observation (x, y)). The reader should note that the first 3 terms of (13) are the Dirac phase (10) along two perpendicular segments that continuously connect the initial point (x_0, y_0) to the point of observation (x, y), *in a clockwise sense* (see for example the red-arrow paths in Fig.1(b)). But apart from this Dirac phase, we also have nonlocal contributions from B_z and its flux within the "observation rectangle" (see i.e. the rectangle

being formed by the red- and green-arrow paths in Fig.1(b)). Below we will directly verify that (13) is indeed a solution of (9) (even for $B_z(x',y') \neq 0$ for $(x',y') \neq (x,y)$), i.e. of the system of PDEs

$$\frac{\partial \Lambda(x,y)}{\partial x} = A_x(x,y) \qquad and \qquad \frac{\partial \Lambda(x,y)}{\partial y} = A_y(x,y). \tag{14}$$

(Although the former is trivially satisfied (at least for cases where interchanges of integrals with derivatives are legitimate), for the latter to be verified one needs to simply substitute $\frac{\partial A_x(x',y)}{\partial y}$ with $\frac{\partial A_y(x',y)}{\partial x'} - B_z(x',y)$ and then carry out the integration with respect to x' – the reader should note the crucial appearance (and proper placement) of $\mathbf{x_0}$ in (13) for the verification of both (14)). It should be noted again that (13) satisfies (14) even for nonzero B_z (i.e. when the particle passes through unequal magnetic fields in remote regions), in contradistinction to the standard result (10). (For the benefit of the reader we clearly provide in the next Section all the steps for the direct verification of (13)).

Equivalently, we will later obtain the result

$$\Lambda(x,y) = \Lambda(x_0,y_0) + \int_{x_0}^{x} A_x(x',\mathbf{y_0})dx' + \int_{y_0}^{y} A_y(x,y')dy' + \left\{ -\int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x',y') + h(y) \right\} \tag{15}$$

$$with \;\; h(y) \;\; chosen \;\; so \;\; that \quad \left\{ -\int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x',y') + h(y) \right\} : \text{ is independent of } \; y,$$

and again the reader should note that, apart from the first 3 terms (the Dirac phase (10) along the two other (alternative) perpendicular segments (connecting (x_0,y_0) to (x,y)), now *in a counterclockwise sense* (the green-arrow paths in Fig.1(b))), we also have nonlocal contributions from the flux of B_z that is enclosed within the same "observation rectangle" (that is naturally defined by the four segments of the two solutions (Fig.1(b))). It can also be easily verified that (15) also satisfies the system (14) (for this $\frac{\partial A_y(x,y')}{\partial x}$ needs to be substituted with $\frac{\partial A_x(x,y')}{\partial y'} + B_z(x,y')$ and then integration with respect to y' needs to be carried out, with the proper appearance (and placement) of $\mathbf{y_0}$ in (15) now being the crucial element – see direct verification in the next Section).

In all the above, A_x and A_y are the Cartesian components of $\boldsymbol{A}(\mathbf{r}) = \boldsymbol{A}(x,y) = \boldsymbol{A_2}(\mathbf{r}) - \boldsymbol{A_1}(\mathbf{r})$, and, as already mentioned, B_z is the difference between (perpendicu-

lar) magnetic fields that the two systems may experience in regions that *do not contain* the observation point (x, y) (i.e. $B_z(x', y') = (\boldsymbol{B_2}(x', y') - \boldsymbol{B_1}(x', y'))_z = \frac{\partial A_y(x',y')}{\partial x'} - \frac{\partial A_x(x',y')}{\partial y'}$, and, although at the point of observation (x, y) we have $B_z(x, y) = 0$ (already emphasized in the Introductory Sections), this $B_z(x', y')$ can be nonzero for $(x', y') \neq (x, y)$). It should be noted that it is because of $B_z(x, y) = 0$ that the functions $g(x)$ and $h(y)$ of (13) and (15) can be found, and the new solutions therefore exist (and are nontrivial). For the impatient reader, simple physical examples with the associated analytical forms of $g(x)$ and $h(y)$ are given in detail in Section VI.

In the present and following Section we place the emphasis in pointing out (and proving) the new solutions (that apparently have been widely overlooked in the literature). In later Sections, we will see that these results actually demonstrate that the passage of particles through magnetic fields has the effect of cancelling Aharonov-Bohm types of phases. And in the special case of narrow wavepacket states in semiclassical motion we will provide an understanding of this cancellation in terms of the experimentally observed compatibility (or consistency) between the Aharonov-Bohm fringe-displacement and the trajectory-deflection due to the Lorentz force. (The corresponding "electric analog" of this consistency of semi-classical trajectory-behavior will also be pointed out $-$ through an elementary physical picture that is here given for the first time, to the best of our knowledge). However, the above cancellations are true even for completely delocalized states (and the deeper reason for this will be obvious from the derivation of the above two solutions (presented in detail in Appendix D) $-$ the origin of the cancellations being essentially the single-valuedness of phases for simple-connected space). Therefore, these generalized results go beyond the usual Aharonov-Bohm behaviors reviewed in the Introductory Sections, and give an extended description of physical systems in more involved physical arrangements (where the particle also passes through remote fields). [It is also simply added here that cancellations of the above type will be extended and generalized further to cases that also involve the time variable t; these will be presented in later Sections, with a detailed mathematical derivation given in Appendix G. Interpreted in a different way, such cancellations $-$ through the new nonlocal terms $-$ will take away the "mystery" of why certain classical arguments (based on past history and the Faraday's law of Induction) seem to "work" (give the correct Aharonov-Bohm phases in static arrangements, by invoking the *history* of how the experimental set up was built at earlier times). Although we will give very general methods of deriving even more

generalized results, applicable to a large number of physical cases, we will mostly restrict attention to detailed applications that provide a natural remedy for earlier discussed "paradoxes" in time-dependent Aharonov-Bohm configurations, and are indicative of an even more general causal propagation of wavefunction phases in the Schrödinger Picture].

V. ELEMENTARY VERIFICATION OF ABOVE SOLUTIONS (EVEN FOR CASES WITH $B_z \neq 0$ IN REMOTE REGIONS)

In static cases, and simple-connected space, let us call our solution (13) Λ_1, namely

$$\Lambda_1(x,y) = \Lambda_1(x_0,y_0) + \int_{x_0}^{x} A_x(x',y)dx' + \int_{y_0}^{y} A_y(x_0,y')dy' + \left\{ \int_{y_0}^{y} dy' \int_{x_0}^{x} dx' B_z(x',y') + g(x) \right\}$$

with $g(x)$ chosen so that $\left\{ \int^{y}\int^{x} B_z + g(x) \right\}$ is independent of x.

Verification that it solves the system of PDEs (14) (even for $B_z(x',y') \neq 0$ for $(x',y') \neq (x,y)$)**:**

A) $\frac{\partial \Lambda_1(x,y)}{\partial x} = A_x(x,y)$ satisfied trivially ✓

(because {...} is independent of x).

B) $\frac{\partial \Lambda_1(x,y)}{\partial y} = \int_{x_0}^{x} \frac{\partial A_x(x',y)}{\partial y} dx' + A_y(x_0,y) + \int_{x_0}^{x} B_z(x',y)dx' + \frac{\partial g(x)}{\partial y}$,

(the last term being trivially zero, $\frac{\partial g(x)}{\partial y} = 0$), and then with the substitution

$\frac{\partial A_x(x',y)}{\partial y} = \frac{\partial A_y(x',y)}{\partial x'} - B_z(x',y)$

we obtain

$\frac{\partial \Lambda_1(x,y)}{\partial y} = \int_{x_0}^{x} \frac{\partial A_y(x',y)}{\partial x'} dx' - \int_{x_0}^{x} B_z(x',y)dx' + A_y(x_0,y) + \int_{x_0}^{x} B_z(x',y)dx'$.

(i) We see that the 2nd and 4th terms of the right-hand-side (rhs) *cancel each other*, and (ii) the 1st term of the rhs is $\int_{x_0}^{x} \frac{\partial A_y(x',y)}{\partial x'} dx' = A_y(x,y) - A_y(x_0,y)$.

Hence finally

$\frac{\partial \Lambda_1(x,y)}{\partial y} = A_y(x,y)$. ✓

We have directly shown therefore (by "going backwards") that the basic system of PDEs (14) is indeed satisfied by our **generalized** solution $\Lambda_1(x,y)$, **even for any nonzero**

$B_z(x', y')$ (in regions $(x', y') \neq (x, y)$; recall that always $B_z(x, y) = 0$). To fully appreciate the above simple proof, the reader is urged to look at the cases of "striped" B_z-distributions in the next Section, the point of observation (x, y) always lying outside the strips, so that the above function $g(x)$ can easily be determined, and the new solutions really *exist* - and they are nontrivial. (As already noted, a formal derivation of the above solution (13) - rather than its above "backwards" verification - is given in Appendix D).

In a completely analogous way, one can easily see that our alternative solution (eq.(15)) also satisfies the basic system of PDEs above. Indeed, if we call our second static solution (eq.(15)) Λ_2, namely

$$\Lambda_2(x, y) = \Lambda_2(x_0, y_0) + \int_{x_0}^{x} A_x(x', y_0)dx' + \int_{y_0}^{y} A_y(x, y')dy' + \left\{ -\int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x', y') + h(y) \right\}$$

with $h(y)$ chosen so that $\left\{ -\int^{x}\int^{y} B_z + h(y) \right\}$: is independent of y,

then we have (even for $B_z(x', y') \neq 0$ for $(x', y') \neq (x, y)$):

A) $\frac{\partial \Lambda_2(x,y)}{\partial y} = A_y(x, y)$ satisfied trivially ✓

(because $\{...\}$ is independent of y).

B) $\frac{\partial \Lambda_2(x,y)}{\partial x} = A_x(x, y_0) + \int_{y_0}^{y} \frac{\partial A_y(x,y')}{\partial x} dy' - \int_{y_0}^{y} B_z(x, y')dy' + \frac{\partial h(y)}{\partial x}$,

(the last term being trivially zero, $\frac{\partial h(y)}{\partial x} = 0$), and then with the substitution

$\frac{\partial A_y(x,y')}{\partial x} = \frac{\partial A_x(x,y')}{\partial y'} + B_z(x, y')$

we obtain

$\frac{\partial \Lambda_2(x,y)}{\partial x} = A_x(x, y_0) + \int_{y_0}^{y} \frac{\partial A_x(x,y')}{\partial y'} dy' + \int_{y_0}^{y} B_z(x, y')dy' - \int_{y_0}^{y} B_z(x, y')dy'$.

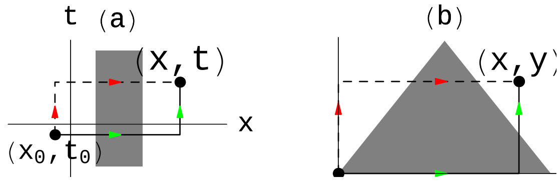
(i) We see that the last two terms of the rhs *cancel each other*, and

(ii) the 2nd term of the rhs is $\int_{y_0}^{y} \frac{\partial A_x(x,y')}{\partial y'} dy' = A_x(x, y) - A_x(x, y_0)$.

Hence finally

$\frac{\partial \Lambda_2(x,y)}{\partial x} = A_x(x, y)$. ✓

Once again, all the above are true for any nonzero $B_z(x', y')$ (in regions $(x', y') \neq (x, y)$).



And a clear understanding of this proof (through the actual analytical form of $h(y)$) is given by the "striped" examples of next Section.

VI. SIMPLE EXAMPLES: NEW RESULTS SHOWN IN EXPLICIT FORM

To see how the above solutions appear in nontrivial cases (and how they give completely new results, i.e. *not differing from the usual ones* (*i.e. from the Dirac phase*) *by a mere constant*) let us first take examples of striped B_z-distributions in space:

(a) For the case of an extended *vertical* strip - parallel to the y-axis, such as in Fig.1(a) (imagine t replaced by y) (i.e. for the case that the particle has actually passed through nonzero B_z, hence through *different* magnetic fields in the two (mapped) systems), then, for x located outside (and on the right of) the strip, the quantity $\int_{y_0}^{y} dy' \int_{x_0}^{x} dx' B_z(x', y')$ in Λ_1 *is already independent of* x (since a displacement of the (x, y)-corner of the rectangle to the right, along the x-direction, does not change the enclosed magnetic flux − see Fig. 1(a) for the analogous (x, t)-case that will be discussed in following Sections). Indeed, in this case the above quantity (the enclosed flux within the "observation rectangle") does not depend on the x-position of the observation point, but on the positioning of the boundaries of the B_z-distribution in the x-direction (better, on the constant width of the strip) − as the x-integral does not give any further contribution when the dummy variable x' goes out of the strip. In fact, in this case the enclosed flux depends on y as we discuss below (but, again, not on x). Hence, for this case, the function $g(x)$ can be easily determined: it can be taken as $g(x) = 0$ (up to a constant C), because then the condition for $g(x)$ stated in the solution (13) (namely, that the quantity in brackets must be independent of x) is indeed satisfied.

We see therefore above that for this setup, the nonlocal term in the solution *survives* (the quantity in brackets is nonvanishing), but *it is not constant*: as already noted, this enclosed flux depends on y (since the enclosed flux *does change* with a displacement of the

(x, y)-corner of the rectangle upwards, along the y-direction, as the y-integral *is* affected by the positioning of y $-$ the higher the positioning of the observation point the more flux is enclosed inside the observation rectangle). Hence, by looking at the alternative solution $\Lambda_2(x, y)$, the quantity $\int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x', y')$ is *dependent on* y, so that $h(y)$ must be chosen as $h(y) = +\int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x', y')$ (up to the same constant C) in order to *cancel* this y-dependence, so that its own condition stated in the solution (15) (namely, that the quantity in brackets must be independent of y) is indeed satisfied**;** as a result, the quantity in brackets in solution Λ_2 disappears and there is no nonlocal contribution in Λ_2 (for $C = 0$). (Of course, if we had used a $C \neq 0$, the nonlocal contributions would be shared between the two solutions in a different manner, but without changing the Physics when we take the *difference* of the two solutions (see below)). [The crucial point in the above is, once again that, because $B_z = 0$ at (x, y), any displacement of this observation point to the right does *not* change the flux enclosed inside the "observation rectangle"**;** and this makes the new solutions (i.e. the functions $g(x)$ and $h(y)$) exist].

With these choices of $h(y)$ and $g(x)$, we already have new results (compared to the standard ones of the integrals of potentials). I.e. one of the two solutions, namely Λ_1 **is** affected nonlocally by the enclosed flux (and this flux is **not** constant). Spelled out clearly, the two results are:

$$\Lambda_1(x, y) = \Lambda_1(x_0, y_0) + \int_{x_0}^{x} A_x(x', y) dx' + \int_{y_0}^{y} A_y(x_0, y') dy' + \int_{y_0}^{y} dy' \int_{x_0}^{x} dx' B_z(x', y') + C$$

$$\Lambda_2(x, y) = \Lambda_2(x_0, y_0) + \int_{x_0}^{x} A_x(x', y_0) dx' + \int_{y_0}^{y} A_y(x, y') dy' + C.$$

And now it is easy to note that, if we subtract the two solutions Λ_1 and Λ_2, the result is *zero* (because the line integrals of the vector potential $\boldsymbol{A}$ in the two solutions are in opposite senses in the (x, y) plane, hence their difference leads to a *closed* line integral of $\boldsymbol{A}$, which is in turn equal to the enclosed magnetic flux, and this flux always happens to be of opposite sign from that of the enclosed flux that explicitly appears as a nonlocal contribution of the B_z-fields (i.e. the term that survives in Λ_1 above)**.** Hennce, the two solutions are *equal.* [We of course everywhere assumed, as usual, single-valuedness of Λ at the initial point (x_0, y_0),

i.e. $\Lambda_1(x_0,y_0)=\Lambda_2(x_0,y_0)$; matters of multivaluedness of Λ at the observation point (x,y) will be addressed later (Section X)].

The reader should probably note that, formally speaking, the above *equality* of the two solutions is due to the fact that the x-independent quantity in brackets of the 1st solution (13) is equal to the function $h(y)$ of the 2nd solution (15), and the y-independent quantity in brackets of the 2nd solution (15) is equal to the function $g(x)$ of the first solution (13). This will turn out to be a general behavioral pattern of the two solutions in simple-connected space, that will be valid for any shape of B_z-distribution, as will be shown in Section X.

This vanishing of $\Lambda_1(x,y)-\Lambda_2(x,y)$ is a cancellation effect that is emphasized further (and generally proved) later below (and can be viewed as a generalization of the Werner & Brill experimental observations[11] to general delocalized states, as will be fully discussed, in completely *physical terms*, in Section X). It basically originates from the single-valuedness of Λ at (x,y) for simple-connected space. This effect is generalized even further in later Sections (i.e. also to cases of combined spacetime variables x,y,t) for the van Kampen thought-experiment[12] (where we will have a combination of spatial multiple-connectivity at an initial instant t_0, and simple-connectivity in (x,t) and (y,t) planes).

(b) In the "dual case" of an extended *horizontal* strip - parallel to the x-axis, the proper choices (for y above the strip) are basically reverse (i.e. we can now take $h(y)=0$ and $g(x)=-\int_{y_0}^{y}dy'\int_{x_0}^{x}dx'B_z(x',y')$ (since the flux enclosed in the rectangle now depends on x, but not on y), with both choices always up to a common constant) and once again we can easily see, upon subtraction of the two solutions, a similar cancellation effect. In this case as well, the results are again new (a nonlocal term survives now in Λ_2). Again spelled out clearly, these are:

$$\Lambda_1(x,y)=\Lambda_1(x_0,y_0)+\int_{x_0}^{x}A_x(x',y)dx'+\int_{y_0}^{y}A_y(x_0,y')dy'+C$$

$$\Lambda_2(x,y)=\Lambda_2(x_0,y_0)+\int_{x_0}^{x}A_x(x',y_0)dx'+\int_{y_0}^{y}A_y(x,y')dy'-\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y')+C$$

(their difference also being zero – a generalized Werner & Brill cancellation (see Section X for further discussion)). Again here the crucial point is that, because the B_z-configuration

does *not contain the point* (x, y), a displacement of this observation point upwards does *not* change the flux inside the "observation rectangle"; this makes the new solutions (i.e the functions $g(x)$ and $h(y)$) exist.

(c) If we want cases that are more involved (i.e. with the nonlocal contributions appearing nontrivially in **both** solutions Λ_1 and Λ_2 and with $g(x)$ and $h(y)$ not being "immediately visible"), we must consider different shapes of B_z-distributions. One such case is a triangular one that is shown in Fig.1(b) (for simplicity an equilateral triangle, and with the initial point $(x_0, y_0) = (0, 0)$) and with the point of observation (x, y) being fairly close to the triangle's right side as in the Figure. Note that for such a configuration, the part of the magnetic flux that is inside the "observation rectangle" (defined by the right upper corner (x, y)) depends on **both** x **and** y. It turns out, however, that this (x and y)−dependent enclosed flux can be written as a sum of separate x- and y-contributions, so that appropriate $g(x)$ and $h(y)$ can still be found (each one of them must be chosen so that it only cancels the corresponding variable's dependence of the enclosed flux). For a homogeneous B_z it is a rather straightforward exercise to determine this enclosed part, i.e. the common area between the observation rectangle and the equilateral triangle, and from this we can find the appropriate $g(x)$ that will cancel the x-dependence, and the appropriate $h(y)$ that will cancel the y-dependence. These appropriate choices turn out to be

$$g(x) = B_z \left[-(\sqrt{3}ax - \frac{\sqrt{3}}{2}x^2) + \frac{\sqrt{3}}{4}a^2 \right] + C \tag{16}$$

and

$$h(y) = B_z \left[(ay - \frac{y^2}{\sqrt{3}}) - \frac{\sqrt{3}}{4}a^2 \right] + C \tag{17}$$

with a being the side of the equilateral triangle. (We again note that a physical arbitrariness described by the common constant C, does not play any role when we take the difference of the two solutions (13) and (15)). We should emphasize that expressions (16) and (17), if combined with (13) or (15), give the nontrivial nonlocal contributions of the difference B_z of the remote magnetic fields on Λ of each solution (hence on the phase of the wavefunction of each wavepacket travelling along each path) at the observation point (x, y), that always lies outside the B_z-triangle. (We mention again that in the case of completely spread-out

states, the equality of the two solutions at the observation point essentially demonstrates the uniqueness (single-valuedness) of the phase in simple-connected space). Further physical discussion of the above cancellations, and a semiclassical interpretation, is given later in Section X and in the final Sections of the paper.

Finally, in more "difficult" geometries, i.e. when the shape of the B_z-distribution is such that the enclosed flux does *not* decouple in a sum of separate x- and y-contributions, such as cases of circularly shaped B_z-distributions, it is advantageous to solve the system (9) directly in non-Cartesian (i.e. polar) coordinates. The results of such a procedure in polar coordinates are given in Appendix E (see eqs (71)-(74)). A general comment that can be made for general shapes is that, depending on the geometry of shape of the B_z-distribution, an appropriate change of variables (to a new coordinate system) may first be needed, so that generalized solutions of the system (9) can be found (namely, so that the enclosed flux inside the *transformed* observation rectangle (i.e. a slice of an annular section in the case of polar coordinates) can be written as a sum of separate (transformed) variables), and then the same methodology (as in the above Cartesian cases) can be followed.

Finally, the reader who may wonder how the usual Aharonov-Bohm result comes out from the present formulation (that contains the additional nonlocal terms), must first read Section X (where the most general solutions (24) and (25) for this 2-D static case are derived, containing additional "multiplicities") and then Appendix F that gives a detailed derivation of the standard Aharonov-Bohm results.

VII. EXAMPLE OF GENERALIZED SOLUTIONS IN DYNAMICAL CASES (WITH ELECTRIC FIELDS, EVEN FOR CASES WITH $\boldsymbol{E \neq 0}$ IN REMOTE SPACETIME REGIONS)

Let us now look at a case with full time-dependence. Although it may be possible to guess the corresponding generalized results, i.e. for a spatially-one-dimensional (x, t)-problem (by appropriate Euclidian rotation of the above solutions in 4-D space), let us nevertheless start from the beginning and give a full physical discussion − as this *is* the case that actually led us to the above generalized solutions, and a case associated with a number of misleading arguments (and often incorrect results) propagating in the literature.

Let us then first focus on the simplest case of one-dimensional quantum systems, i.e. a

single quantum particle of charge q, but in the presence of the most general (spatially nonuniform and time-dependent) vector and scalar potentials, and ask the following question: what is the gauge function $\Lambda(x,t)$ that takes us from (maps) a system with potentials $A_1(x,t)$ and $\phi_1(x,t)$ to a system with potentials $A_2(x,t)$ and $\phi_2(x,t)$ (meaning the usual mapping (7) between the wavefunctions of the two systems through the phase factor $\frac{q}{\hbar c}\Lambda(x,t)$)? [Once again we should keep in mind that for this mapping to be possible we *must* assume that at the point (x,t) of observation (or "measurement" of Λ or the wavefunction Ψ) we have equal electric fields ($E_i = -\nabla\phi_i - \frac{1}{c}\frac{\partial A_i}{\partial t}$), namely

$$-\frac{\partial\phi_2(x,t)}{\partial x} - \frac{1}{c}\frac{\partial A_2(x,t)}{\partial t} = -\frac{\partial\phi_1(x,t)}{\partial x} - \frac{1}{c}\frac{\partial A_1(x,t)}{\partial t} \tag{18}$$

(so that the A's and ϕ's in (18) can indeed satisfy the basic system of equations (8), or equivalently, of the system of equations (21) below – as can be seen by taking the $\frac{1}{c}\frac{\partial}{\partial t}$ of the 1st and the $\frac{\partial}{\partial x}$ of the 2nd of the system (21) and adding them together). But again, we will *not* exclude the possibility of the two systems passing through *different* electric fields in different regions of spacetime, i.e. for $(x',t') \neq (x,t)$. In fact, this possibility **will come out naturally** from a careful solution of the basic system (21); it is for example straightforward for the reader to immediately verify that the results (22) or (23) that will be derived below (and will contain contributions of electric field-differences from remote regions of spacetime) indeed satisfy the basic input system of equations (21), something that will be explicitly verified in the next Section].

Returning to the question on the appropriate Λ that takes us from the set (A_1,ϕ_1) to the set (A_2,ϕ_2), we note that, in cases of static vector potentials ($A(x)$'s) *and* spatially uniform scalar potentials ($\phi(t)$'s) the answer usually given is the well-known

$$\Lambda(x,t) = \Lambda(x_0,t_0) + \int_{x_0}^{x} A(x')dx' - c\int_{t_0}^{t}\phi(t')dt' \tag{19}$$

with $A(x) = A_2(x) - A_1(x)$ and $\phi(t) = \phi_2(t) - \phi_1(t)$ (and it can be viewed as a combination of (10) and (12) (or of (3) and (4)), being immediately applicable to the description of cases of *combined* magnetic and electric Aharonov-Bohm effects reviewed in the Introductory Sections).

In the most general case (and with the variables x and t being **completely uncorrelated**), it is often stated in the literature [as in eq. (57) of Ref.[4], taken for $\mathbf{v} = 0$, a very

good example to point to, since that article does not use a path-integral language, but a canonical formulation with uncorrelated variables] that the appropriate Λ has a form that is a plausible extention of (19), namely

$$\Lambda(x,t) = \Lambda(x_0,t_0) + \int_{x_0}^{x} [A_2(x',t) - A_1(x',t)]\, dx' - c\int_{t_0}^{t} [\phi_2(x,t') - \phi_1(x,t')]\, dt', \tag{20}$$

and as already pointed out in Section II, this form is certainly *incorrect* for uncorrelated variables x and t (the reader can easily verify that the system of equations (21) below is *not* satisfied by (20) $-$ see again Section II if needed, especially the paragraph after eq.(6)). We will find in the present work that the correct form consists of two major modifications: (i) The first leads to the natural appearance of a *path* that continuously connects initial and final points in spacetime, a property that (20) *does not have* [indeed, if the integration curves of (20) are drawn in the (x,t)-plane, they do *not* form a continuous path from (x_0,t_0) to (x,t)]. The reader can immediately see eqs.(22) and (23) that will be given below for the corrected "path-forms" in the line integrals of potentials (and these are represented by the red-arrow and green-arrow paths of Fig.1(a)). (ii) And the second modification is highly nontrivial: it consists of nonlocal contributions of classical electric field-differences from remote regions of spacetime. We will discuss below the consequences of these terms and we will later show that such nonlocal contributions also appear (in an extended form) in more general situations, i.e. they are also present in higher spatial dimensionality (and they then also involve remote magnetic fields in combination with the electric ones); these lead to modifications of ordinary Aharonov-Bohm behaviors or have other important consequences, one of them being a natural remedy of Causality "paradoxes" in time-dependent Aharonov-Bohm experiments.

The form (20) commonly used is of course motivated by the well-known Wu & Yang[13] nonintegrable phase factor, that has a phase equal to $\int A_\mu dx^\mu = \int A dx - c\int \phi dt$, a form that appears naturally within the framework of path-integral treatments, or generally in physical situations where narrow wavepackets are implicitly assumed for the quantum particle: the integrals appearing in (20) are then taken along particle trajectories (hence spatial and temporal variables *not* being uncorrelated, but being connected in a particular manner $x(t)$ to produce the path; all integrals are therefore basically only time-integrals). But even then, eq.(20) is valid only when these trajectories are always (in time) and everywhere (in space)

inside identical classical fields for the two (mapped) systems. Here, however, we will be focusing on what a canonical (and not a path-integral or other semiclassical) treatment leads to; this will cover the general case of arbitrary wavefunctions that can even be completely spread-out in space, and will also allow the particle to travel through different electric fields for the two systems in remote spacetime regions (e.g. $E_2(x,t') \neq E_1(x,t)$ if $t' < t$ etc.).

It is therefore clear that in order to find the appropriate $\Lambda(x,t)$ that answers the question posed above in full generality will require a careful solution of the system of PDEs (8), applied to only one spatial variable, namely

$$\frac{\partial \Lambda(x,t)}{\partial x} = A(x,t) \qquad and \qquad -\frac{1}{c}\frac{\partial \Lambda(x,t)}{\partial t} = \phi(x,t) \tag{21}$$

(with $A(x,t) = A_2(x,t) - A_1(x,t)$ and $\phi(x,t) = \phi_2(x,t) - \phi_1(x,t)$). This system of PDEs is solved in detail in Appendix B, and leads to two alternative solutions, one being

$$\Lambda(x,t) = \Lambda(x_0,t_0) + \int_{x_0}^{x} A(x',t)dx' - c\int_{t_0}^{t} \phi(x_0,t')dt' + \left\{ c\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' E(x',t') + g(x) \right\} + \tau(t_0) \tag{22}$$

with $g(x)$ chosen so that the quantity $\left\{ c\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' E(x',t') + g(x) \right\}$ is independent of x, and (from an inverted route of integrations) the other solution being

$$\Lambda(x,t) = \Lambda(x_0,t_0) + \int_{x_0}^{x} A(x',t_0)dx' - c\int_{t_0}^{t} \phi(x,t')\, dt' + \left\{ -c\int_{x_0}^{x} dx' \int_{t_0}^{t} dt' E(x',t') + \hat{g}(t) \right\} + \chi(x_0) \tag{23}$$

with $\hat{g}(t)$ chosen so that the quantity $\left\{ -c\int_{x_0}^{x} dx' \int_{t_0}^{t} dt' E(x',t') + \hat{g}(t) \right\}$ is independent of t.

In the above $E = (E_2 - E_1)$ is the difference of perpendicular electric fields in the two systems, which can be nonvanishing at regions remote to the observation point (x,t) (see below). (Note again that at the point of observation $E(x,t) = 0$, signifying the basic fact that the fields in the two systems are identical at the point of observation (x,t)). Solutions (22) and (23) can be viewed as the (formal) analogs of (13) and (15) correspondingly, although they hide in them much richer Physics because of their dynamic character (see Section IX).

(The additional constant last terms will be shown in Section IX to be related to possible multiplicities of Λ, and they are zero in simple-connected spacetimes). Also note again that the integrations of potentials in (22) and (23) indeed form paths that continuously connect (x_0, t_0) to (x, t) in the xt-plane (the red-arrow and green-arrow paths of Fig.1(a)), a property that the incorrectly used solution (20) does *not* have.

The reader is once again provided with the direct verification that (22) or (23) are indeed solutions of the basic system of PDEs (21) in the Section that follows.

VIII. VERIFICATION OF SOLUTIONS AND SIMPLE DYNAMICAL PHYSICAL EXAMPLES

Let us call our first solution (eq.(22)) for simple-connected spacetime Λ_3, namely

$$\Lambda_3(x,t) = \Lambda_3(x_0,t_0) + \int_{x_0}^{x} A(x',t)dx' - c\int_{t_0}^{t} \phi(x_0,t')dt' + \left\{ c\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' E(x',t') + g(x) \right\}$$

with $g(x)$ chosen so that $\left\{ c\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' E(x',t') + g(x) \right\}$ is independent of x.

Verification that it solves the system of PDEs (21) (even for $E(x',t') \neq 0$ for $(x',t') \neq (x,t)$)**:**

A) $\frac{\partial \Lambda_3(x,t)}{\partial x} = A(x,t)$ satisfied trivially ✓

(because $\{....\}$ is independent of x).

B) $-\frac{1}{c}\frac{\partial \Lambda_3(x,t)}{\partial t} = -\frac{1}{c}\int_{x_0}^{x} \frac{\partial A(x',t)}{\partial t}dx' + \phi(x_0,t) - \int_{x_0}^{x} E(x',t)dx' - \frac{1}{c}\frac{\partial g(x)}{\partial t}$,

(the last term being trivially zero, $\frac{\partial g(x)}{\partial t} = 0$), and then with the substitution
$-\frac{1}{c}\frac{\partial A(x',t)}{\partial t} = \frac{\partial \phi(x',t)}{\partial x'} + E(x',t)$
we obtain
$-\frac{1}{c}\frac{\partial \Lambda_3(x,t)}{\partial t} = \int_{x_0}^{x} \frac{\partial \phi(x',t)}{\partial x'}dx' + \int_{x_0}^{x} E(x',t)dx' + \phi(x_0,t) - \int_{x_0}^{x} E(x',t)dx'.$

(i) We see that the 2nd and 4th terms of the rhs *cancel each other*, and

(ii) the 1st term of the rhs is $\int\limits_{x_0}^{x} \frac{\partial \phi(x',t)}{\partial x'} dx' = \phi(x,t) - \phi(x_0,t)$.

Hence finally

$-\frac{1}{c}\frac{\partial \Lambda_3(x,t)}{\partial t} = \phi(x,t)$. ✓

We have directly shown therefore that the basic system of PDEs (21) is indeed satisfied by our **generalized** solution $\Lambda_3(x,t)$, **even for any nonzero** $E(x',t')$ (in regions $(x',t') \neq (x,t)$). (The reader is again reminded that always $E(x,t) = 0$). Once again, the function $g(x)$ owes its existence to the fact that the spacetime point of observation (x,t) is outside the E-distribution (hence the term *nonlocal,* used for the effect of the field-difference E on the phases), and the reader can clearly see this in the "striped" E-distributions of the examples that follow later in this Section.

In a completely analogous way, one can easily see that our alternative solution (eq.(23)) also satisfies the basic system of PDEs above. Indeed, if we call our second (alternative) solution (eq.(23)) for simple-connected spacetime Λ_4, namely

$$\Lambda_4(x,t) = \Lambda_4(x_0,t_0) + \int\limits_{x_0}^{x} A(x',t_0)dx' - c\int_{t_0}^{t} \phi(x,t')\,dt' + \left\{ -c\int\limits_{x_0}^{x} dx' \int\limits_{t_0}^{t} dt' E(x',t') + \hat{g}(t) \right\}$$

with $\hat{g}(t)$ chosen so that $\left\{ -c\int\limits_{x_0}^{x} dx' \int\limits_{t_0}^{t} dt' E(x',t') + \hat{g}(t) \right\}$ is independent of t,

then we have (even for $E(x',t') \neq 0$ for $(x',t') \neq (x,t)$)**:**

A) $-\frac{1}{c}\frac{\partial \Lambda_4(x,t)}{\partial t} = \phi(x,t)$ satisfied trivially ✓

(because $\{....\}$ is independent of t).

B) $\frac{\partial \Lambda_4(x,t)}{\partial x} = A(x,t_0) - c\int\limits_{t_0}^{t} \frac{\partial \phi(x,t')}{\partial x} dt' - c\int\limits_{t_0}^{t} E(x,t')dt' + \frac{\partial \hat{g}(t)}{\partial x}$,

(the last term being trivially zero, $\frac{\partial \hat{g}(t)}{\partial x} = 0$), and then with the substitution

$\frac{\partial \phi(x,t')}{\partial x} = -E(x,t') - \frac{1}{c}\frac{\partial A(x,t')}{\partial t'}$

we obtain

$\frac{\partial \Lambda_4(x,t)}{\partial x} = A(x,t_0) + c\int\limits_{t_0}^{t} E(x,t')dt' + \int\limits_{t_0}^{t} \frac{\partial A(x,t')}{\partial t'} dt' - c\int\limits_{t_0}^{t} E(x,t')dt'$.

(i) We see that the 2nd and 4th terms of the rhs *cancel each other*, and

(ii) the 3rd term of the rhs is $\int\limits_{t_0}^{t} \frac{\partial A(x,t')}{\partial t'} dt' = A(x,t) - A(x,t_0)$.

Hence finally

$\frac{\partial \Lambda_4(x,t)}{\partial x} = A(x,t)$. ✓

Once again, all the above are true for any nonzero $E(x',t')$ (in regions $(x',t') \neq (x,t)$).

To see again how the above solutions appear in nontrivial cases (and how they give new results, i.e. not differing from the usual ones by a mere constant) let us take analogous examples of strips as earlier, but now in spacetime:

(a) For the case of the extended *vertical* strip (parallel to the t-axis) of Fig.1(a) (the case of a one-dimensional capacitor that is (arbitrarily and variably) charged for all time), then, for x located outside (and on the right of) the capacitor, the quantity $c\int\limits_{t_0}^{t} dt' \int\limits_{x_0}^{x} dx' E(x',t')$ in Λ_3 is *already independent of* x (since a displacement of the (x,t)-corner of the rectangle to the right, along the x-direction, does not change the enclosed "electric flux", see Fig.1(a))**;** hence in this case the function $g(x)$ can be taken as $g(x) = 0$ (up to a constant C), because then the condition for $g(x)$ stated in the solution (22) (namely, that the quantity in brackets must be independent of x) is indeed satisfied. (Note again that the above x-independence of the enclosed "electric flux" is important for the existence of $g(x)$).

So for this setup, the nonlocal term in the solution *survives* (the quantity in brackets is nonvanishing), but *it is not constant***:** this enclosed flux depends on t (since the enclosed flux *does change* with a displacement of the (x,t)-corner of the rectangle upwards, along the t-direction). Hence, by looking at the alternative solution $\Lambda_4(x,t)$, the quantity $c\int\limits_{x_0}^{x} dx' \int\limits_{t_0}^{t} dt' E(x',t')$ is dependent on t, so that $\hat{g}(t)$ must be chosen as $\hat{g}(t) = +c\int\limits_{x_0}^{x} dx' \int\limits_{t_0}^{t} dt' E(x',t')$ (up to the same constant C) in order to *cancel* this t-dependence, so that its own condition stated in the solution (23) (namely, that the quantity in brackets must be independent of t) is indeed satisfied**;** as a result, the quantity in brackets in solution Λ_4 disappears and there is no nonlocal contribution in Λ_4 (for $C = 0$). (Once again, if we had used a $C \neq 0$, the nonlocal contributions would be differently shared between the

two solutions, but without changing the Physics when we take the *difference* of the two solutions). [The reader should once again note the crucial fact that the point of observation (x,t) is outside the E-distribution, which makes the existence of functions $g(x)$ and $\hat{g}(t)$ possible].

With these choices of $\hat{g}(t)$ and $g(x)$, we already have new results (compared to the standard ones of the integrals of potentials). I.e. one of the two solutions, namely Λ_3 **is** affected nonlocally by the enclosed flux (and this flux is **not** constant). Spelled out clearly, the two results are:

$$\Lambda_3(x,t) = \Lambda_3(x_0,t_0) + \int_{x_0}^{x} A(x',t)dx' - c\int_{t_0}^{t} \phi(x_0,t')dt' + c\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' E(x',t') + C$$

$$\Lambda_4(x,t) = \Lambda_4(x_0,t_0) + \int_{x_0}^{x} A(x',t_0)dx' - c\int_{t_0}^{t} \phi\left(x,t'\right) dt' + C$$

(and their difference, as mentioned above, is zero - denoting what might be called a generalized Werner & Brill cancellation in spacetime).

(b) In the "dual case" of an extended *horizontal* strip - parallel to the x-axis (that corresponds to a nonzero electric field in all space that has however a finite duration T), the proper choices (for observation time instant $t > T$) are basically reverse (i.e. we can now take $\hat{g}(t) = 0$ and $g(x) = -c\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' E(x',t')$ (since the "electric flux" enclosed in the "observation rectangle" now depends on x, but not on t), with both choices always up to a common constant) and once again we can easily see, upon subtraction of the two solutions, a similar cancellation effect. In this case again, the results are also new (a nonlocal term survives now in Λ_4). Again spelled out clearly, these are:

$$\Lambda_3(x,t) = \Lambda_3(x_0,t_0) + \int_{x_0}^{x} A(x',t)dx' - c\int_{t_0}^{t} \phi(x_0,t')dt' + C$$

$$\Lambda_4(x,t) = \Lambda_4(x_0,t_0) + \int_{x_0}^{x} A(x',t_0)dx' - c\int_{t_0}^{t} \phi\left(x,t'\right) dt' - c\int_{x_0}^{x} dx' \int_{t_0}^{t} dt' E(x',t') + C$$

(their difference also being zero – a generalized Werner & Brill cancellation in spacetime).

(c) And again, if we want cases that are more involved (with the nonlocal contributions appearing nontrivially in **both** solutions Λ_3 and Λ_4 and with $g(x)$ and $\hat{g}(t)$ not being "immediately visible") we must again consider different shapes of E-distribution. One such case (the triangular) was already shown in Fig.1(b) (for the magnetic case, which however is completely analogous). For such a triangular case the choices of $g(x)$ and $\hat{g}(t)$ will be different from the above and this will result in different roles of the nonlocal terms (and these nontrivial results, or more accurately, their analogs for the magnetic case, were given earlier in closed analytical form, eqs (16) and (17)). [And even cases of curved shapes can be addressed more generally (when the shape is such that the "flux" does *not* decouple in a sum of separate spatial and temporal contributions), i.e. by solving the basic system of PDEs directly in polar coordinates (the results being analogous to the ones given in Appendix E for the magnetic case, see eqs (71)-(74))].

The reader should note again that, in all the above examples in simple-connected spacetime, the x-independent quantity in brackets of the 1st solution (22) is equal to the function $\hat{g}(t)$ of the 2nd solution (23), and the t-independent quantity in brackets of the 2nd solution (23) is equal to the function $g(x)$ of the 1st solution (22). This mathematical pattern is what leads to the above mentioned cancellations, and it is generally proven (i.e. for any form of E-distribution in the (x, t)-plane) in the Section that follows.

IX. COMMENTS ON THE GENERAL BEHAVIOR OF THE (x, t)-SOLUTIONS

Let us first summarize (and prove in generality) some of the behavioral patterns that we saw in the above examples and then continue on other properties (i.e. an account of multiplicities of Λ in multiple-connected spacetimes that we left out, which are described by the constants $\tau(t_0)$ and $\chi(x_0)$). First, in (22) or (23) note the proper appearance and placement of x_0 and t_0 that gives a "path-sense" to the line integrals of potentials in each solution (with the path consisting of two straight and perpendicular line segments, continuously connecting the initial point (x_0, t_0) to the final point (x, t) for each solution). And there are naturally two possible paths of this type that connect the initial point (x_0, t_0) with the final point (x, t) (the solution (22) having a clockwise and the solution (23) having a counterclockwise sense, as in Fig.1(a)); with this construction a natural *observation rectangle* is then formed

(see Fig. 1(a)), within which the enclosed "electric fluxes" (in spacetime) appear to be crucial (showing up as nonlocal terms of contributions of the electric field difference (recall that $E(x',t') = E_2(x',t') - E_1(x',t')$) from regions of time and space that are remote to the observation point (x,t)). The appearance of these nonlocal terms (of the electric field difference) in $\Lambda(x,t)$ from regions of spacetime (x',t') far from the observation point (x,t) seems to have a direct effect on the wavefunction phases at (x,t) (through the phase mapping that connects the two quantum systems). The actual manner in which this happens is of course determined by the form of functions $g(x)$ or $\hat{g}(t)$ (the existence of which lies in the fact that the spacetime observation point (x,t) is always outside the E-distributions): these functions must be chosen in such a way that they satisfy their respective conditions, as these are stated after (22) or (23) respectively. We saw, for example, that if we have a distribution of E in the (x,t)-plane in the form of an extended *strip* parallel to the t-axis, the function $g(x)$ can be taken as $g(x) = 0$ (up to a constant C), and that $\hat{g}(t)$ must be chosen as $\hat{g}(t) = +c\int_{x_0}^{x} dx' \int_{t_0}^{t} dt' E(x',t')$ (up to the same constant C) in order to *cancel* the t-dependence of the enclosed "flux". Furthermore, with these choices of $\hat{g}(t)$ and $g(x)$, it is easy to see that, if we subtract the two solutions (22) and (23), the result is *zero* (because the line integrals of potentials A and ϕ in the two solutions are in opposite senses in the (x,t) plane, hence their difference leads to a *closed* line integral, which is in turn equal to the enclosed electric flux, and this flux always happens to be of opposite sign from that of the enclosed flux that explicitly appears as a nonlocal contribution of the E-fields (i.e. the term that survives in Λ_3 in case (a) of last Section). Such cancellation effects in dynamical cases are important and will be discussed (and generalized) further in Section XII.

Let us however give here a general proof of the above cancellations (i.e. for any variable dependence of E-distribution). First, by looking at the general structure of solutions (22) and (23), we note that in both forms, the last constant terms ($\tau(t_0)$ and $\chi(x_0)$) are only present in cases where Λ is expected to be multivalued (this comes from the definitions of $\tau(t_0)$ and $\chi(x_0)$, and is shown in Appendix C, see eqs (63) and (64)) and therefore these constant quantities are nonvanishing in cases of motion only in multiple-connected spacetimes (leading to phenomena of the electric Aharonov-Bohm type (see the analogous discussion in Appendix F, on the easier-to-follow magnetic case)). In such multiple-connected cases these last terms turn out to be simply equal (in absolute value) to the enclosed fluxes in regions of spacetime

that are physically inaccessible to the particle (in the electric Aharonov-Bohm setup, for example, it turns out that $\tau(t_0) = -\chi(x_0) =$ enclosed "electric flux" in spacetime). Although such cases can also be covered by our method below, let us for the moment ignore them (set them to zero) and focus on cases of motion in simple-connected spacetimes. Then the two solutions (22) and (23) are actually *equal* for *any* E-distribution. This is rigorously shown in Appendix C. [It is also shown there that the x-independent (hence t-dependent) quantity in brackets of the 1st solution (22) is equal to the function $\hat{g}(t)$ of the 2nd solution (23) − and the t-independent (hence x-dependent) quantity in brackets of the 2nd solution (23) is equal to the function $g(x)$ of the 1st solution (22). Because of this, it is straightforward to see (by subtracting the two solutions) the mathematical reason for the occurence of the cancellations noted earlier, for *any* shape of E-distribution].

In spite therefore of the simplicity of the above considered 1-D system, we are already in a position to draw certain very general conclusions on the possible physical consequences of the new nonlocal terms of the electric fields appearing in the solutions (22) and (23). One can immediately see from the above considerations (or from the formal proof of Appendix C) that these temporally-nonlocal contributions have the tendency of cancelling the contributions from the A- and ϕ-integrals. This already gives an indication of cancellations that might also occur in cases of higher spatial dimensionality (where line-integrals of A's, for example, can be related to enclosed *magnetic* fluxes). This *is* actually the case in the van Kampen thought-experiment that will be discussed later in Section XII − although the cancellations there will be slightly more delicate, actually involving a balance among 3 variables, and with the actual *senses* of spatial closed line-integrals in the (x, y)-plane being nontrivially important. (Moreover, *instead of lying outside of simple strips, the spatial point* (x, y) *will in that case lie outside a light-cone*, leading to results that are *causal*, as we shall see).

Finally, with respect to $\tau(t_0)$ and $\chi(x_0)$, we show in the same Appendix C their already noted properties**:** ordinarily (in simple-connectivity) they are zero, or in the most general case (of multiple-connectivity) they are related to physically inaccessible enclosed fluxes. [We should also note here that the case of the electric Aharonov-Bohm setup, with the particles traveling inside distinct equipotential cages with scalar potentials that last for a finite duration, is the prototype of *multiple-connectivity in spacetime*, a fact first noted by Iddings and reported by Noerdlinger[14]. We will see later (Section XII) that this feature is *not* present in the van Kampen thought-experiment, hence an electric Aharonov-Bohm

argument should not really be invoked in that case (as van Kampen did) because of this lack of multiple-connectedness in spacetime].

Before, however, leaving this simple (x,t)-case, we should finally emphasize that this (or any other) contribution of electric fields is *not* present at the level of the basic Lagrangian, and the view holds in the literature (see e.g. [15]) that, because of this absence, electric fields cannot contribute *directly* to the phase of the wavefunctions. This conclusion originates from the path-integral approach (that is almost always followed), but, nevertheless, our present work shows that fields *do* contribute nonlocally. A more general discussion on this issue is given in the final Section, after discussion of the van Kampen thought-experiment, and also in connection to related path-integral approaches[16].

X. AGAIN ON THE (x,y)-MAGNETIC CASE

After having discussed fully the simple (x,t)-case, let us for completeness give the analogous (Euclidian-rotated in 4-D spacetime) derivation for (x,y)-variables and briefly discuss the properties of the simpler static solutions, but now in full generality (also including possible multi-valuedness of Λ in the usual magnetic Aharonov-Bohm cases). We will simply need to apply the same methodology (of solution of a system of PDEs) to such static spatially two-dimensional cases (so that now different (remote) magnetic fields for the two systems, perpendicular to the 2-D space, will arise). For such cases we need to solve the system of PDEs already shown in (14), namely

$$\frac{\partial\Lambda(x,y)}{\partial x} = A_x(x,y) \qquad \textit{and} \qquad \frac{\partial\Lambda(x,y)}{\partial y} = A_y(x,y).$$

By following then the procedure described in detail in Appendix D, we finally obtain the following general solution

$$\Lambda(x,y) = \Lambda(x_0,y_0)+\int_{x_0}^{x} A_x(x',y)dx'+\int_{y_0}^{y} A_y(x_0,y')dy'+\left\{\int_{y_0}^{y} dy'\int_{x_0}^{x} dx' B_z(x',y') + g(x)\right\}+f(y_0) \tag{24}$$

$$with \;\; g(x) \;\; chosen \;\; so \;\; that \;\; \left\{ \int\limits_{y_0}^{y} dy' \int\limits_{x_0}^{x} dx' B_z(x', y') + g(x) \right\} : \text{ is independent of } \; x,$$

which is basically the example shown earlier in (13) but with included multiplicities through the extra constant $f(y_0)$ (that for simple-connected space can be set to zero). The result (24) applies to cases where the particle passes through *different* magnetic fields (recall that $B_z = (\boldsymbol{B}_2 - \boldsymbol{B}_1)_z$) in spatial regions that are remote to (i.e. do not contain) the observation point (x, y). Alternatively, by following the reverse route of integrations (see Appendix D), we finally obtain the following alternative general solution

$$\Lambda(x, y) = \Lambda(x_0, y_0) + \int_{x_0}^{x} A_x(x', y_0) dx' + \int_{y_0}^{y} A_y(x, y') dy' + \left\{ -\int\limits_{x_0}^{x} dx' \int\limits_{y_0}^{y} dy' B_z(x', y') + h(y) \right\} + \hat{h}(x_0) \tag{25}$$

$$with \;\; h(y) \;\; chosen \;\; so \;\; that \;\; \left\{ -\int\limits_{x_0}^{x} dx' \int\limits_{y_0}^{y} dy' B_z(x', y') + h(y) \right\} : \text{ is independent of } \; y,$$

which is basically the example shown earlier in (15) but with included multiplicities through the extra constant $\hat{h}(x_0)$. One can actually show that the two solutions are equivalent (i.e. (13) and (15) for a simple-connected region are equal[17]), a fact that can be proven in a way similar to the (x, t)-cases of Section IX (the actual proof being given in Appendix C). (For the case of multiple-connectivity of the two-dimensional space, a discussion of the actual values of the multiplicities $f(y_0)$ and $\hat{h}(x_0)$ is given later in this Section, with the proofs presented in Appendix F).

As we saw in the examples of Section VI, in case of a striped-distribution of the magnetic field difference B_z, the functions $g(x)$ and $h(y)$ in (24) and (25) (or equivalently in (13) and (15)) have to be chosen in ways that are compatible with their corresponding constraints (stated after (24) and (25)) and are completely analogous to the above discussed (x, t)-cases. (In all cases, the fact that the observation point (x, y) is always outside the nonzero-B_z regions is crucial for the existence of these functions). By then taking the *difference* of (13) and (15) we obtain that the "Aharonov-Bohm phase" (the one originating from the *closed* line integral of A's) is exactly cancelled by the additional nonlocal term of

the magnetic fields (that the particle passed through). As already mentioned earlier, this is reminiscent of the cancellation of phases (broadly speaking, a cancellation between the "Aharonov-Bohm phase" and the semiclassical phase picked up by the trajectories) observed in the early experiments of Werner & Brill[11] for particles passing through *full* (nonvanishing) magnetic fields, and our method seems to provide a very natural justification: as our results are completely general (and for delocalized states in a simple-connected region they basically describe the single-valuedness of Λ), they are also valid and applicable to cases of narrow wavepackets (or states that describe semiclassical motion) that pass through nonvanishing magnetic fields, which *was* the case of the Werner & Brill experiments. (A similar cancellation of an electric Aharonov-Bohm phase also occurs for particles passing through a static electric field as we saw in Section VIII). We conclude that, for static cases, and when particles pass through fields, the new nonlocal terms reported in the present work lead quite generally to a cancellation of Aharonov-Bohm phases that had earlier been sketchily noticed and only at the semiclassical level.

Since we already mentioned that the deep origin of the above cancellations is the single-valuedness of Λ in simple-connected space, we should add for completeness that the rigorous proof of the uniqueness at each spatial point (single-valuedness) of Λ for completely delocalized states in simple-connected space can be given in a directly analogous way to the proof given in Appendix C for the (x, t)-case of Section IX. What is however more important to point out here is that the above cancellations for semiclassical trajectories (that pass through a nonzero magnetic field) can alternatively be understood as a *compatibility* between the Aharonov-Bohm fringe-displacement and the trajectory-deflection due to the Lorentz force (i.e. the semiclassical phase picked up due to the optical path difference of the two deflected trajectories *exactly cancels* (is *opposite in sign* from) the Aharonov-Bohm phase picked up by the trajectories due to the enclosed flux). [We may mention that this is also related to the well-known overall rigid displacement of the single-slit envelope of the two-slit diffraction pattern, displacement that only occurs if the wavepackets actually pass through a nonzero field (and not in genuine Aharonov-Bohm cases)]. These issues are further discussed in the final Section, where some popular reports in the literature (Feynman[18], Felsager[19], Batelaan & Tonomura[20]) are given a minor correction (of a sign). Similarly, and by also including time t (and by again correcting a sign-error propagating in the standard literature) we will give an explanation of why certain classical arguments (invoking the

past t-dependent history of the experimental set up) seem to be successful (in giving the correct result for a static Aharonov-Bohm phase).

Another point of interest concerning the above found nonlocal contributions of fields is the plausible question of *what shape* the field distributions must have (or more accurately, their part enclosed inside the *observation rectangle*) so that the enclosed flux can be decoupled to a sum of functions of separate variables, in order for the solutions obtained above to exist and be immediately applicable (i.e. for the functions $g(x)$ and $h(y)$ to be directly possible to determine: each of them must then only *partially* cancel the corresponding x *or* y dependence, respectively). We already provided an example of such a distribution of a homogeneous B_z (the triangular one) in Section VI (see the nontrivial results (16) and (17)). And as mentioned in Section VI, in cases of circularly shaped distributions (where the enclosed flux may not be decoupled in x and y terms), it is advantageous to solve the system directly in polar coordinates. By following a similar procedure (of solving the system of PDEs resulting from (9)) in polar coordinates (ρ, φ), namely

$$\frac{\partial\Lambda(\rho,\varphi)}{\partial\rho} = A_\rho(\rho,\varphi) \qquad and \qquad \frac{1}{\rho}\frac{\partial\Lambda(\rho,\varphi)}{\partial\varphi} = A_\varphi(\rho,\varphi)$$

with steps completely analogous to the above described procedure, one can obtain analogs of solutions (24) and (25) given in Appendix E (see eqs (71) and (73)). [In such a case, the observation rectangle has now given its place to a slice of an annular section]. These matters however deserve further investigation, of a more mathematical type, in applications of the above theory to specific shape-geometries.

Finally, for completeness we discuss the issue of multiplicities (the last constant terms of (24) and (25)) in case of spatial multiple-connectivity (such as the standard magnetic Aharonov-Bohm case, in which we can take $g(x) = 0$ *and* $h(y) = 0$, since the enclosed magnetic flux is independent of both x and y). We take up this issue in detail in Appendix F, where it is proven that $\hat{h}(x_0) = - \ f(y_0) =$ enclosed magnetic flux (a constant, independent of x and y). Since $f(y_0)$ cancels out the $\int\limits_{y_0}^{y} dy' \int\limits_{x_0}^{x} dx' B_z(x', y')$ term, and $\hat{h}(x_0)$ cancels out the $-\int\limits_{x_0}^{x} dx' \int\limits_{y_0}^{y} dy' B_z(x', y')$ term, the two solutions are then reduced to the usual solutions of mere A-integrals along the two paths (i.e. the standard Dirac phase, with no nonlocal

contributions).

XI. FULL (x, y, t)-CASE

Finally, let us look at the most general spatially-two-dimensional and time-dependent case. This combines effects of (perpendicular) magnetic fields (which, if present only in physically-inaccessible regions, can have Aharonov-Bohm consequences) with the temporal nonlocalities of electric fields (parallel to the plane) found in previous Sections. By working again in Cartesian spatial coordinates, we now have to deal with the full system of PDEs

$$\frac{\partial\Lambda(x,y,t)}{\partial x} = A_x(x,y,t), \qquad \frac{\partial\Lambda(x,y,t)}{\partial y} = A_y(x,y,t), \qquad -\frac{1}{c}\frac{\partial\Lambda(x,y,t)}{\partial t} = \phi\,(x,y,t)\,. \tag{26}$$

This exercise is considerably longer than the previous ones but important to solve, in order to see in what manner the solutions of this system are able to *combine* the spatial and temporal nonlocal effects found above. There are now 3!=6 alternative integration routes to follow for solving this system (and, in addition to this, the results in intermediate steps tend to proliferate). The corresponding (rather long) procedure for solving the system (26) is described in detail in Appendix G, and 2 out of the 12 solutions that can be derived turn out to be the most crucial for the discussion that will follow in the next Section. First, by following steps similar to the above, the following temporal generalization of (25) is obtained

$$\Lambda(x,y,t) = \Lambda(x_0,y_0,t) + \int_{x_0}^{x} A_x(x',y_0,t)dx' + \int_{y_0}^{y} A_y(x,y',t)dy' +$$
$$+\left\{-\int\limits_{x_0}^{x} dx' \int\limits_{y_0}^{y} dy' B_z(x',y',t) + G(y,t)\right\} + f(x_0,t) \tag{27}$$

$$\textit{with}\;\; G(y,t)\;\; \textit{such that} \quad \left\{-\int\limits_{x_0}^{x} dx' \int\limits_{y_0}^{y} dy' B_z(x',y',t) + G(y,t)\right\}\textbf{:}\quad \textsf{is independent of}\;\; y,$$

and from this point on, the third equation of the system (26) is getting involved to determine the nontrivial effect of scalar potentials on $G(y,t)$**.** Indeed, by combining it with (27) there results a wealth of patterns, one of them leading finally to our first solution

$$\Lambda(x,y,t)=\Lambda(x_0,y_0,t_0)+\int_{x_0}^{x}A_x(x',y_0,t)dx'+\int_{y_0}^{y}A_y(x,y',t)dy'-\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y',t)+G(y,t_0)-$$

$$-c\int_{t_0}^{t}\phi(x_0,y_0,t')dt'+c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E_x(x',y,t')+c\int_{t_0}^{t}dt'\int_{y_0}^{y}dy'E_y(x_0,y',t')+F(x,y)+f(x_0,t_0) \tag{28}$$

with the functions $G(y,t_0)$ and $F(x,y)$ to be chosen in such a way as to satisfy the following 3 independent conditions:

$$\left\{G(y,t_0)-\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y',t_0)\right\}:\ is\ \textsf{independent of}\ \ y, \tag{29}$$

which is of course a special case of the condition on $G(y,t)$ above (see after (27)) applied at $t=t_0$, and the other 2 turn out to be of the form

$$\left\{F(x,y)+c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E_x(x',y,t')\right\}:\ is\ \textsf{independent of}\ \ x, \tag{30}$$

$$\left\{F(x,y)+c\int_{t_0}^{t}dt'\int_{y_0}^{y}dy'E_y(x,y',t')\right\}:\ is\ \textsf{independent of}\ \ y. \tag{31}$$

It is probably important to inform the reader that for the above results the Faraday's law is crucial (see Appendix G). As for the constant quantity $f(x_0,t_0)$ appearing in (28), it again describes possible effects of multiple-connectivity at the instant t_0 (which are absent for simple-connected spacetimes, but will be crucial in the discussion of the van Kampen thought-experiment to be discussed in the next Section).

Eq. (28) is our first solution. It is now crucial to note that an alternative form of solution (with the functions $G's$ and F satisfying the *same* conditions as above) can be derived (see Appendix G) which turns out to be

$$\Lambda(x,y,t)=\Lambda(x_0,y_0,t_0)+\int_{x_0}^{x}A_x(x',y_0,t)dx'+\int_{y_0}^{y}A_y(x,y',t)dy'-\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y',t_0)+G(y,t_0)-$$

$$
-c\int\limits_{t_0}^{t}\phi(x_0,y_0,t')dt' + c\int\limits_{t_0}^{t}dt'\int\limits_{x_0}^{x}dx'E_x(x',y_0,t') + c\int\limits_{t_0}^{t}dt'\int\limits_{y_0}^{y}dy'E_y(x,y',t') + F(x,y) + f(x_0,t_0).
\tag{32}
$$

In this alternative solution we note that, in comparison with (28), the line-integrals of $\boldsymbol{E}$ have changed to the *other* alternative "path" (note the difference in the placement of the coordinates of the initial point (x_0, y_0) in the arguments of E_x and E_y compared to (28)) and they happen to have the same sense as the $\boldsymbol{A}$-integrals, while simultaneously the magnetic flux difference shows up with its value at the initial time t_0 rather than at t. This alternative form will be shown to be useful in cases where we want to directly compare physical situations in the present (at time t) and in the past (at time t_0), and the above noted change of sense of $\boldsymbol{E}$-integrals (compared to (28)) will be crucial in the discussion that follows in the next Section. (It is also important here to note that, in the form (32), the electric fields have already incorporated the effect of radiated B_z-fields in space (through the Maxwell's equations, see Appendix G), and this is why at the end only the B_z at t_0 appears explicitly).

Once again the reader can directly verify that (28) or (32) indeed satisfy the basic input system (26). (This verification is considerably more tedious than the earlier ones but rather straightforward).

But a last mathematical step remains**:** in order to discuss the van Kampen case, namely an enclosed (and physically inaccessible) magnetic flux (which however is *time-dependent*), it is important to have the analogous forms through a reverse route of integrations (see Appendix G), where at the end we will have the reverse "path" of $\boldsymbol{A}$-integrals (so that by taking the *difference* of the resulting solution and the above solution (28) (or (32)) will lead to the *closed* line integral of $\boldsymbol{A}$ which will be immediately related to the van Kampen's magnetic flux (at the instant t)). By following then the reverse route, and by applying a similar strategy at every intermediate step, we finally obtain the following solution (the spatially "dual" of (28)), namely

$$
\Lambda(x,y,t) = \Lambda(x_0,y_0,t_0) + \int_{x_0}^{x} A_x(x',y,t)dx' + \int_{y_0}^{y} A_y(x_0,y',t)dy' + \int\limits_{x_0}^{x}dx'\int\limits_{y_0}^{y}dy'B_z(x',y',t) + \hat{G}(x,t_0) -
$$

$$
-c\int_{t_0}^{t}\phi(x_0,y_0,t')dt' + c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E_x(x',y_0,t') + c\int_{t_0}^{t}dt'\int_{y_0}^{y}dy'E_y(x,y',t') + F(x,y) + \hat{h}(y_0,t_0) \tag{33}
$$

with the functions $\hat{G}(x,t_0)$ and $F(x,y)$ to be chosen in such a way as to satisfy the following 3 independent conditions:

$$
\left\{\hat{G}(x,t_0) + \int_{y_0}^{y}dy'\int_{x_0}^{x}dx'B_z(x',y',t_0)\right\} : \ is \ \text{independent of} \ \ x, \tag{34}
$$

$$
\left\{F(x,y) + c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E_x(x',y,t')\right\} : \ is \ \text{independent of} \ \ x, \tag{35}
$$

$$
\left\{F(x,y) + c\int_{t_0}^{t}dt'\int_{y_0}^{y}dy'E_y(x,y',t')\right\} : \ is \ \text{independent of} \ \ y, \tag{36}
$$

where again for the above results the Faraday's law is crucial (see Appendix G). The corresponding analog of the alternative form (32) (where B_z appears at t_0) is more important (for the discussion of the next Section) and turns out to be

$$
\Lambda(x,y,t) = \Lambda(x_0,y_0,t_0) + \int_{x_0}^{x}A_x(x',y,t)dx' + \int_{y_0}^{y}A_y(x_0,y',t)dy' + \int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y',t_0) + \hat{G}(x,t_0) -
$$

$$
-c\int_{t_0}^{t}\phi(x_0,y_0,t')dt' + c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E_x(x',y,t') + c\int_{t_0}^{t}dt'\int_{y_0}^{y}dy'E_y(x_0,y',t') + F(x,y) + \hat{h}(y_0,t_0) \tag{37}
$$

with $\hat{G}(x,t_0)$ and $F(x,y)$ following the same 3 conditions above. The constant term $\hat{h}(y_0,t_0)$ again describes possible multiplicities at the instant t_0; it is absent for simple-connected spacetimes, but will be crucial in the discussion of the van Kampen thought-experiment.

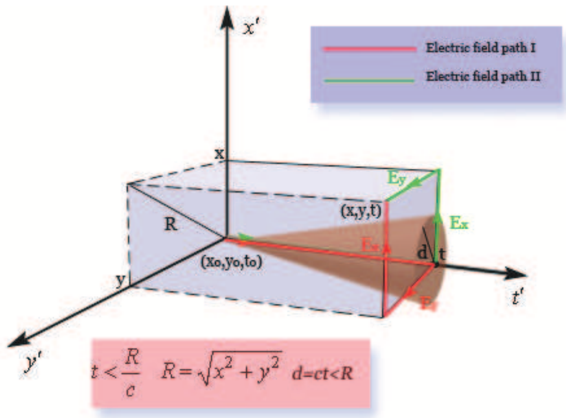
In (33) (and in (37)), note the "alternative paths" (compared to solution (28) (and (32))) of line integrals of $\boldsymbol{A}$'s (or of $\boldsymbol{E}$'s). But the most crucial element for what follows is the need to *exclusively* use the forms (32) and (37) (where B_z only appears at t_0), and the fact that,

within each solution, the sense of $\boldsymbol{A}$-integrals is the *same* as the sense of the $\boldsymbol{E}$-integrals. (This is *not* true in the other solutions where $B_z(..,t)$ appears, as the reader can directly see). These facts will be crucial to the discussion that follows, which briefly addresses the so called "van Kampen paradox".

XII. THE VAN KAMPEN THOUGHT-EXPERIMENT – HOW THE ABOVE SOLUTIONS ENFORCE CAUSALITY

In an early work[12] van Kampen considered a genuine Aharonov-Bohm case, with a magnetic flux (physically inaccessible to the particle) which, however, is time-dependent**:** van Kampen envisaged turning on the flux very late, or equivalently, observing the interference of the two wavepackets (on a distant screen) very early, earlier than the time it takes light to travel the distance to the screen, hence using the (instantaneous nature of the) Aharonov-Bohm phase to transmit information (on the existence of a confined magnetic flux somewhere in space) *superluminally.* Indeed, the Aharonov-Bohm phase at any later instant t is determined by differences of $\frac{q}{\hbar c}\Lambda(\mathbf{r},t)$, with $\Lambda(\mathbf{r},t) = \int_{\mathbf{r}_0}^{\mathbf{r}} \boldsymbol{A}(\mathbf{r}',t)\boldsymbol{.}d\mathbf{r}' +$ *const.* (which basically results as a special case (but in higher dimensionality) of the incorrect expression (20) (or (5)) in the temporal gauge $\phi = 0$, the constant being $\Lambda(\mathbf{r}_0,t_0)$). However, let us for this case utilize instead our results (32) and (37) above, where we have the additional appearance of the nonlocal E-terms (and of the B_z-term at t_0).

In order to be slightly more general than van Kampen, let us for example assume that the inaccessible magnetic flux had the value $\Phi(t_0)$ at t_0, and then it started changing with time. By using a narrow wavepacket picture like van Kampen, we can then subtract (32) and (37) in order to find the phase difference at a time t that is smaller than the time required for light to reach the observation point (x,y) (i.e. $t < \frac{L}{c}$, with L the corresponding distance). For a spatially-confined magnetic flux $\Phi(t)$, the functions G, $\hat{G}$ and F in the above solutions can then all be taken zero**:** **(i)** their conditions are all satisfied for a flux $\Phi(t)$ that is not spatially-extended (hence, from (29) and (34) we obtain $G = \hat{G} = 0$ since the integrals in brackets are all independent of x and y), and **(ii)** for $t < \frac{L}{c}$, the integrals of E_x and E_y in conditions (30) and (31) (or in (35) and (36)) are already independent of *both* x and y (since $E_x(x,y,t') = E_y(x,y,t') = 0$ for all $t' < t < \frac{L}{c}$, with (x,y) the observation point (since at instant t, the $\boldsymbol{E}$-field has not yet reached the spatial point (x,y) of the screen), and



therefore all integrations of E_x and E_y with respect to x' and y' will be contributing only up to a light-cone (see Fig.2) and they will therefore give results that are *independent of the integration upper limits x and y* – basically a generalization of the striped cases that we saw earlier but now to the case of 3 spatio-temporal variables (with now the spatial point (x, y) being outside the light-cone defined by t (see Fig.2; in this Figure the initial spatial point (x_0, y_0), taken for simplicity at $(0, 0)$, has been supposed to be in the area of the inaccessible flux $\Phi(t)$, so that, for $R = \sqrt{(x-x_0)^2 + (y-y_0)^2}$, we have indeed that $R \sim L$, hence that ct (which is $< L$, see above) is also $< R$, as written on the Figure))**;** we therefore rigorously obtain $F = 0$). Moreover, the multiplicities (f and $\hat{h}$) lead to cancellation of the B_z-terms (at t_0) in exactly the same manner as outlined in the static case earlier (at the end of Section X). By choosing then the temporal gauge $\phi = 0$ like van Kampen, we have for the difference (32) $-$ (37) at the point and instant of observation the following result

$$\Delta\Lambda(x,y,t) = \int_{x_0}^{x} A_x(x', y_0, t)dx' + \int_{y_0}^{y} A_y(x, y', t)dy' - \int_{x_0}^{x} A_x(x', y, t)dx' - \int_{y_0}^{y} A_y(x_0, y', t)dy' +$$

$$+ c\int_{t_0}^{t} dt' \left\{ \int_{x_0}^{x} dx' E_x(x', y_0, t') + \int_{y_0}^{y} dy' E_y(x, y', t') - \int_{x_0}^{x} dx' E_x(x', y, t') - \int_{y_0}^{y} dy' E_y(x_0, y', t') \right\}. \tag{38}$$

In (38) the sum of the four A-integrals gives the *closed* line-integral of vector $\boldsymbol{A}$ around the *observation rectangle* at time t (in the positive sense) and it is equal to the instantaneous magnetic flux $\Phi(t)$ (that leads to the "usual" magnetic Aharonov-Bohm phase)**;** the sum of the four E-integrals inside the brackets in the last terms (originating from our nonlocal

contributions) gives the *closed* line-integral of vector $\boldsymbol{E}$ around the same rectangle at any arbitrary t', and in the same (positive) sense (something we wouldn't have if we had taken the first type of solutions, (28) and (33) $-$ this signifying the importance of taking the right form, the one that contains B_z at t_0 (with the t-propagation of B_z *having already been incorporated* in the E_x and E_y terms of (32) and (37))). By denoting therefore the closed loop integral (around the rectangle) as $\oint$ always in the positive sense (and with the understanding that the rectangle's upper right corner is the spatial point of observation (x, y)), (38) reads

$$\Delta\Lambda(x,y,t) = \oint \mathbf{A}(\mathbf{r}',t)\boldsymbol{.}d\mathbf{r}' + c\int_{t_0}^{t} dt' \oint \mathbf{E}(\mathbf{r}',t')\boldsymbol{.}d\mathbf{r}' \tag{39}$$

which, with $\oint \mathbf{A}(\mathbf{r}',t)\boldsymbol{.}d\mathbf{r}' = \Phi(t)$ the instantaneous enclosed magnetic flux and with the help of Faraday's law $\oint \mathbf{E}(\mathbf{r}',t')\boldsymbol{.}d\mathbf{r}' = -\frac{1}{c}\frac{d\Phi(t')}{dt'}$, gives

$$\Delta\Lambda(x,y,t) = \Phi(t) - (\Phi(t) - \Phi(t_0)) = \Phi(t_0). \tag{40}$$

Although $\Delta\Lambda$ is generally t-dependent, we obtain the intuitive (causal) result that, for $t < \frac{L}{c}$ (i.e. if the physical information has not yet reached the screen), the phase-difference turns out to be t-independent, and leads to the magnetic Aharonov-Bohm phase that we *would* observe at t_0. The new nonlocal terms have conspired in such a way as to *exactly cancel* the Causality-violating Aharonov-Bohm phase (that would be proportional to the instantaneous $\Phi(t)$).

This gives an honest resolution of the "van Kampen paradox" within a canonical formulation, without using any vague electric Aharonov-Bohm effect argument as was done by van Kampen (since in the gauge chosen ($\phi = 0$) there are no scalar potentials – and, most importantly, *there is no multiple-connectivity in* (x, t)*-plane* as in the electric Aharonov-Bohm case[14]). In this van Kampen thought experiment the particle actually passes through the electric and magnetic fields that are radiated outside the confined magnetic flux, and the electric type of phase that recovers Causality is actually an example of our nonlocal terms. The recovery of Causality is the result of the action of these new nonlocal terms, in a type of "generalized Werner & Brill cancellation in spacetime" (the earlier strips having now given their place to a light-cone). An additional physical element (in comparison to van Kampen's

electric phase interpretation) is that, for the above cancellation, it is not only the E-fields but also the t-propagation in space of the B_z-fields (the full "radiation field") that plays a role.

Finally, a number of other forms of solutions can be obtained that result from different ordering of integrations of the system (26) (a full list of 12 different (but quite long) results is available, directly verifiable that they satisfy the system (26)). The reader can follow the strategies of solution suggested here and derive the forms that are appropriate to particular physical cases of interest that may be different from the above magnetic case, some potential candidates being the "electric analog" of the van Kampen thought-experiment, or its *bound state analog* in nanorings. For the former we pay particular attention below; for the latter, and especially for 1-D nanorings (or other nanoscopic devices) driven by a t-dependent magnetic flux, the new nonlocal terms are expected to be of relevance if they are included in standard treatments[21], and the effects are expected to appear in the PetaHertz range. (Similarly we might expect a nontrivial role in cases of quantal astrophysical objects due to the large distances involved (hence retardation effects being more pronounced)).

For the "electric analog" of the van Kampen case, we note that, although this has never really been discussed in the literature (in such terminology), nevertheless, it has been essentially briefly mentioned in Appendix B of Peshkin[8] (where the point is made about what happens when *first* the particle exits the cages, and *only then* we switch on the outside electric field, together with the comment of the author that the results must be "consistent with ordinary ideas about Causality"; Peshkin correctly states: "One cannot wait for the electron to pass and only later switch on the field to cause a physical effect"). As our new nonlocal terms seem to be especially suited for addressing such Causality issues, let us slightly expand on this point: in this most authoritative (and carefully written) review of the Aharonov-Bohm effect in the literature, Peshkin uses (for the electric effect) a solution-form (his eq.(B.5) together with (B.6)) based on (20), i.e. the "standard result" (but applied to a spatially-dependent scalar potential) $-$ but he clearly states that it is an approximation (and actually later in the review, he states that this form cannot be a solution for all t). Indeed, from the present work we learn that Peshkin's eqs (B.5) and (B.6) do *not* give the solution when the scalar potential depends on spatial variables (because the spatial variables inside the potential will give $-$ through its nonzero gradient $-$ an extra vector potential (that will result from $\nabla\Lambda$), hence an extra minimal substitution in the Hamiltonian H, violating

therefore the mapping between two pre-determined systems that we want to achieve). As we saw in the present work, the correct solution for all t and in all space consists of additional nonlocal terms of the appropriate form. If we view the form (B.5) and (B.6) of ref. [8] as an *ansatz*, then it is understandable why a *condition* (Peshkin's eq.(B.8), and later (B.9)) needs to be *enforced* on the electric field outside the cages (in order for the extra (annoying) terms (that show up from expansion of the squared minimal substitution) to vanish and for (B.5) to be a solution). And then Peshkin notes that the extra condition cannot always be satisfied $-$ *it must fail* for some times (hence (B.5) is not really the solution for all times), drawing from this a correct conclusion, namely that "the electron must traverse some region where the electric field *has been*" (earlier). However, the causal feature pointed out above, although mentioned in words, is not dealt with quantitatively. From our present work, it turns out that the total "radiation field" outside the cages is crucial in recovering Causality, in a similar way as in the case presented above in this Section for the usual (magnetic) version of the van Kampen experiment. In this "electric analog" that we are discussing now, the causally-offending part of the electric Aharonov-Bohm phase difference will be cancelled by a magnetic type of phase, that originates from the magnetic field that is associated with the t-dependence of the electric field $\boldsymbol{E}$ outside the cages.

It should be re-emphasized that the correct quantitative physical behavior of the above system for all times comes out from the treatment shown in detail in the present work, with no enforced constraints, but with conditions that come out naturally from the solution of the PDEs. The results that are derived from this careful procedure give the full solutions (correct for all space and for all t)**:** Peshkin's ansatz (B.6) turns out (from an honest and careful solution of the full PDEs) to be augmented by nonlocal terms of the electric fields, and these directly influence the phases of wavefunctions (by always respecting Causality, with no need of enforced conditions) – and can even include the contributions of vector potentials and magnetic fields (through nonlocal magnetic terms in space) associated with the t-variation of the electric field outside the cages, that Peshkin has omitted (as he actually admits in the beginning of his Appendix). As already mentioned, the total "radiation field" outside the cages is crucial in recovering Causality, in a way similar to what was presented above for the usual (magnetic) version of the van Kampen experiment. We conclude that our (exact) results accomplish precisely what Peshkin has in mind in his discussion (on Causality), but in a direct and fully quantitative manner, and with *no ansatz* based on an

incorrect form.

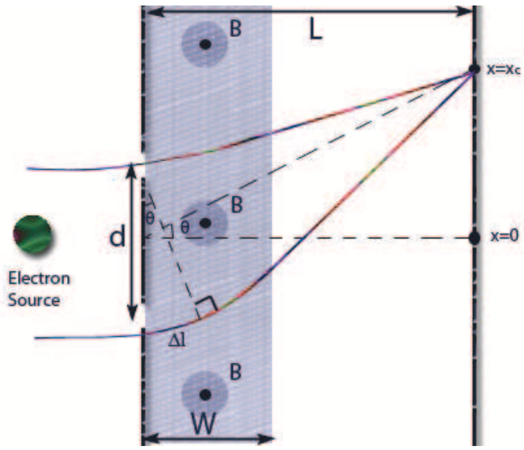
XIII. PHYSICAL DISCUSSION

In attempting to evaluate in a broader sense the crucial nonlocal influences found in all the above physical examples, we should probably first reemphasize that at the level of the basic Lagrangian $L(\mathbf{r},\mathbf{v},t)=\frac{1}{2}m\mathbf{v}^2+\frac{q}{c}\mathbf{v}.\mathbf{A}(\mathbf{r},t)-q\phi(\mathbf{r},t)$ there are no fields present, and the view holds in the literature[15] that electric or magnetic fields cannot contribute *directly* to the phase of quantum wavefunctions. This view originates from the path-integral treatments widely used (where the Lagrangian determines directly the phases of Propagators), but, nevertheless, our canonical formulation treatment shows that fields *do* contribute nonlocally, and they are actually crucial in recovering Relativistic Causality. Moreover, path-integral discussions[16] of the van Kampen case use wave (retarded)-solutions for the vector potentials $\mathbf{A}$ (hence they are treated in Lorenz gauge, which is not sufficiently general**:** even if $\mathbf{A}$ has not yet reached the screen, we can always add a constant $\mathbf{A}$ (a pure gauge) over all space, and there are no more retarded wave-solutions for the potentials, the proposed path-integral resolution of the paradox[16] being, therefore, at least incomplete). Our results are gauge-invariant and take advantage of only the retardation of *fields* $\mathbf{E}$ and $\mathbf{B}$ (true in *any* gauge), and *not* of potentials. In addition, Troudet[16] clearly (and correctly) states that his treatment is good for not highly-delocalized states in space, and that in case of delocalization the proper treatment "would be much more complicated, and would require a much more complete analysis". It is clear that we have provided such a complete analysis in the present work. It should be added that in a recent Compendium of Quantum Physics[22], the "van Kampen paradox" still seems to be thought of as remarkable. It is fair to state that the present work has provided a natural and general resolution, and most importantly, through nonlocal and Relativistically causal propagation of wavefunction phases in the Schrödinger Picture (this point being expanded further below, at the end of the article).

At several places in this paper we have pointed out a number of "misconceptions" in the literature (mostly on the uncritical use of the (standard) Dirac phases even for t-dependent vector potentials and spatially-dependent scalar potentials, which is plainly incorrect for uncorrelated variables), and we have explicitly provided their "healing" through appropriate nonlocal field-terms. It should be re-emphasized here that this is not a merely marginal

misconception, but it appears all over the place in the literature (due to the Feynman path-integral bias); it is even stated by Feynman himself in volume II of his *Lectures on Physics*[18], namely that the simple phase factor $\int^x \boldsymbol{A} \cdot d\mathbf{r}' - c\int^t \phi dt'$ is valid generally, i.e. even for t-dependent fields. Similarly, this erroneous generalization is also explicitly stated in the review on Aharonov-Bohm effects of Erlichson[23] that has given a very balanced view of earlier controversy, and also elsewhere − the books of Silverman[24] being the clearest case that we are aware of with a careful wording about (20) being only restrictedly valid (for t-independent $\boldsymbol{A}$'s and $\mathbf{r}$-independent ϕ's), although even there the nonlocal terms have been missed. We believe that the above misconceptions (and the overlooking of the nonlocal terms) are the basic reason why "it appears that no exact theoretical treatment has been given" (for the electric Aharonov-Bohm effect), as correctly stated by Peshkin in his Appendix B of Ref.[8].

And let us now come to a second type of misconception, that has appeared only in semiclassical conditions − but is essential to mention here, as it is another example that exhibits the merits of our approach (and the deeper physical understanding that our results can lead to). What we learn from the generalized Werner & Brill cancellations pointed out rather emphatically in the present work is that, at the point of observation, the nonlocal terms of classical remote fields have the tendency to contribute a phase *of opposite sign* to the "Aharonov-Bohm phase" (of potentials). We want to point out to the reader that, for semiclassical trajectories, this is actually descriptive of the compatibility (or consistency) of the Aharonov-Bohm fringe-displacement and the associated trajectory-deflection due to the classical (Lorentz) forces. Let us for example look at Fig.15-8 of Feynman[25], or at Fig.2.16 of Felsager[19], where, classical trajectories are deflected after they pass through a strip of a homogeneous magnetic field that is placed on the right of a standard double-slit experimental apparatus (see also our own Fig.3). Both authors determine the semiclassical phase picked up by the trajectories (that have been deflected by the Lorentz force) and they find that they are consistent with the Aharonov-Bohm phase (picked up due to the flux enclosed by the same trajectories). However, it is not very difficult to see that the two phases *have opposite sign* (they are *not equal* as implied by the authors). The reader is also invited to carry out a similar exercise, with particles passing through an analogous homogeneous *electric* field on the right of the double-slit apparatus, with the field being *parallel* to the screen and being switched on for a finite duration T : it then turns out



again that the semiclassical phase picked up is *opposite* to the electric Aharonov-Bohm type of phase (in case this is not immediately clear, a quantitative discussion is given further below). Similarly, in the recent review of Batelaan & Tonomura[20], their Fig.2 contains visual information that is very relevant to our discussion: it is a quite descriptive picture of the wavefronts associated to the classical trajectories, where the authors state that "the phase shift calculated in terms of the Lorentz force is the same as that predicted by the Aharonov-Bohm effect in terms of the vector potential A circling the magnetic bar". The reader, however, should notice once more that the sign of the classical phase-difference is really opposite to the sign of the Aharonov-Bohm phase. The phases are *not equal* as stated, but opposite (see below for a detailed proof). All the above examples may be viewed as a manifestation of the cancellations that have been derived in the present work (for general quantum states), but here they are just special cases for semiclassical trajectories. (We could also restate here that these cancellations have to do with the known rigid displacement of the "single-slit envelope" of the two-slit diffraction pattern in a double-slit experiment, when the particle actually passes through an additional strip of a magnetic field that has been placed on the right of the apparatus).

In case that the reader does not easily see the signs of the relevant phase differences, we provide below elementary proofs of the opposite signs argued above for semiclassical trajectories (and for the special case of small deflections, as usually done in elementary discussions of the standard double-slit setup). First, our Fig.3 demonstrates, by way of an example, the spatial point where the new position of the *central fringe* is now located (after the classical trajectories have been deflected by the additional magnetic strip B). It is shown below why the two phases (semiclassical and Aharonov-Bohm) must indeed be opposite (not

equal) so that, in this new fringe position, the *total* phase difference (i.e. the sum of the above two) is again zero (as it actually *should* be for the central fringe). Indeed, if d is the distance between the two slits, and W the width of the magnetic strip (assumed to be $W << L$ so that the deflections are very small), we have that the "Aharonov-Bohm phase" enclosed between the two classical trajectories (of a particle of charge q) is

$$\Delta\varphi^{AB} = 2\pi \frac{q}{e} \frac{\Phi}{\Phi_0}, \tag{41}$$

with $\Phi_0 = \frac{hc}{e}$ the flux quantum, and $\Phi \approx BWd$ the enclosed flux between the two trajectories (always for small trajectory-deflections), with the deflection originating from the presence of the magnetic strip B and the associated Lorentz forces. On the other hand, the semiclassical phase generally picked up by a trajectory of length l is $\varphi^{semi} = \frac{2\pi}{\lambda} l$, with $\lambda = \frac{h}{\Pi}$ being the de Broglie wavelength (and Π being the classical kinematic momentum mv, with v the speed of the particle, taken almost constant (as usually done) due to the small deflections). The semiclassical phase difference between the 2 classical trajectories is therefore $\Delta\varphi^{semi} = \frac{2\pi}{\lambda}\Delta l$ (with Δl the difference between the 2 semiclassical paths, which in Fig.3 is $\Delta l \approx d\sin\theta \approx d\frac{x_c}{L}$, with x_c being the (displaced) position of the central fringe on the screen, and L the distance between the slit-plane and the screen (note that in Fig.3 we have electrons (hence $q = -e < 0$), the deflections are therefore upward, and we have considered the semiclassical phase difference between the lower trajectory and the upper trajectory (the lower one has a longer path, hence it picks up a higher phase, hence $\Delta\varphi^{semi} > 0 \qquad \checkmark$))). We have therefore

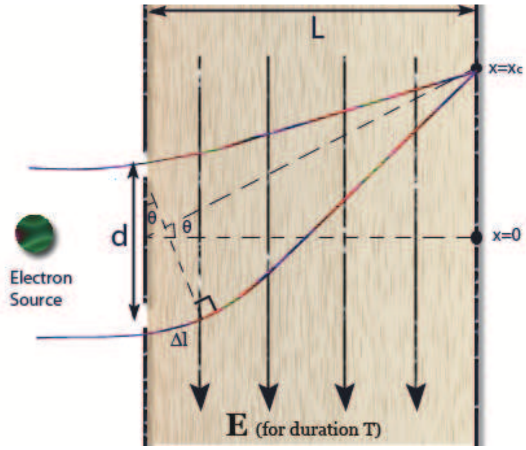
$$\Delta\varphi^{semi} = \frac{2\pi}{\lambda} d \frac{x_c}{L}. \tag{42}$$

Now, the Lorentz force (exerted only during the passage through the thin magnetic strip, hence only during a time interval $\Delta t = \frac{W}{v}$) has a component parallel to the screen (let us call it x-component) that is given by

$$F_x = \frac{q}{c}(\boldsymbol{v} \times \boldsymbol{B})_x = -\frac{q}{c} vB = -\frac{BWq}{c\frac{W}{v}} = -\frac{BWq}{c\Delta t} \tag{43}$$

which shows that there is a change of kinematic momentum (parallel to the screen) $\Delta\Pi_x = -\frac{BWq}{c}$, or, equivalently, a change of parallel speed

$$\Delta v_x = -\frac{BWq}{mc} \tag{44}$$



which is the speed of the central fringe's motion (i.e. its displacement over time along the screen). Although this has been caused by the presence of the thin deflecting magnetic strip, this displacement is occuring uniformly during a time interval $t = \frac{L}{v}$, and this time interval must satisfy

$$\Delta v_x = \frac{x_c}{t} \tag{45}$$

(as, for small displacements, the beams travel most of the time in uniform motion, i.e. $\Delta t << t$). We therefore have that the central fringe displacement must be $x_c = \Delta v_x t = -\frac{BWq}{cm}\frac{L}{v}$, and noting that $mv = \Pi = \frac{h}{\lambda}$, we finally have

$$x_c = -\frac{BWqL\lambda}{hc} \tag{46}$$

(and we note that this displacement is indeed upward (positive) for a negative charge ($q < 0$)). By susbstituting (46) into (42), the lengths L and λ cancel out, and we finally have $\Delta\varphi^{semi} = -2\pi\frac{q}{e}\frac{BWd}{\frac{hc}{e}}$, which with $\frac{hc}{e} = \Phi_0$ the flux quantum, and $BWd \approx \Phi$ the enclosed flux (always for small trajectory-deflections) gives (through comparison with (41)) our final proof that

$$\Delta\varphi^{semi} = -2\pi\frac{q}{e}\frac{\Phi}{\Phi_0} = -\Delta\varphi^{AB}. \tag{47}$$

We note therefore that the semiclassical phase difference (between two trajectories) picked up due to the Lorentz force (exerted on them) is indeed opposite to the Aharonov-Bohm phase due to the magnetic flux enclosed between the same trajectories.

A corresponding electric case is shown in our Fig.4 (discussed here for the first time, to

the best of our knowledge), where now an additional electric field E (pointing downwards) is present everywhere in space, but for only a finite time duration T (which we take to be much shorter than the time of travel $t = \frac{L}{v}$, $T << t$). In this case therefore the electric Lorentz force qE is exerted on the trajectories only during the small time interval $\Delta t = T$ (note the difference with the above magnetic case**;** we now have an electric strip in *time* rather than a magnetic strip in space as we had earlier). Let us then follow an analogous calculation as above but now adopted to this electric case. The electric type of Aharonov-Bohm phase is

$$\Delta\varphi^{AB} = -2\pi \frac{q}{e} \frac{cT\Delta V}{\Phi_0}, \tag{48}$$

with ΔV being the electric potential difference between the two trajectories, hence $\Delta V \approx Ed$ (again for small trajectory-deflections). On the other hand, the semiclassical phase difference between the two trajectories is again given by (42), but the position x_c of the central fringe must now be determined by the electric field force qE **:** The change of kinematic momentum (always parallel to the screen) is now $\Delta\Pi_x = qET$, hence the analog of (44) is now

$$\Delta v_x = \frac{qET}{m} \tag{49}$$

which if combined with (45) (that is obviously valid for this case as well, again for small deflections, due to the $\Delta t = T << t$), and always with $t = \frac{L}{v}$, gives that the central fringe displacement in this case must be $x_c = \Delta v_x t = \frac{qET}{m}\frac{L}{v}$, and using again $mv = \Pi = \frac{h}{\lambda}$, we finally have the following analog of (46)

$$x_c = \frac{qETL\lambda}{h}. \tag{50}$$

(Note again that for a negative charge and a negative electric field (i.e. pointing downwards) the central fringe displacement is indeed upwards). By substituting (50) into (42), the lengths L and λ again cancel out, and we finally have $\Delta\varphi^{semi} = 2\pi d \frac{qETL\lambda}{h} = 2\pi \frac{q}{e} \frac{EdcT}{\frac{hc}{e}}$, which with $\frac{hc}{e} = \Phi_0$ the flux quantum, and through comparison with (48) leads once again to our final proof that

$$\Delta\varphi^{semi} = -\Delta\varphi^{AB}. \tag{51}$$

We note therefore that even in the electric case, the semiclassical phase difference (between two trajectories) picked up due to the Lorentz force (exerted on them) is once again opposite

to the electric Aharonov-Bohm phase due to the electric flux (in spacetime) enclosed between the same trajectories.

We should point out once again, however, that although the above elementary considerations apply to semiclassical motion of narrow wavepackets, in this paper we have given a more general understanding of the above opposite signs that applies to general (even completely delocalized) states, and that originates from our generalized Werner & Brill cancellations.

In a slightly different vein, we should also point out that the above cancellations give a justification of why certain semiclassical arguments that focus on the history of the experimental set up (usually based on Faraday's law for a t-dependent magnetic flux) seem to give at the end a result that is consistent with the result of a static Aharonov-Bohm arrangement. However, there is a again an opposite sign that seems to have been largely unnoticed in such arguments as well (i.e. see the simplest possible argument in Silverman[26], where in his eq.(1.34) there should be an extra minus sign). Our above observation essentially describes the fact that, *if we had actually used* a t-dependent magnetic flux (with its final value being the actual value of our static flux), then the induced electric field (viewed now as a nonlocal term of the present work) would have cancelled the static Aharonov-Bohm phase. Of course now, this t-dependent experimental set up has not been used (the flux is static) and we obtain the usual magnetic Aharonov-Bohm phase, but the above argument (of a "potential experiment" that *could have been carried out*) takes the "mystery" away of why such history-based arguments generally work − although *they have to be corrected with a sign.* The above also gives a rather natural account of the "dynamical nonlocality" character[2] attributed to the various Aharonov-Bohm phenomena (magnetic, electric or combined), although − in the present work − this dynamical quantum nonlocality seems to simultaneously respect Causality, a rather pleasing characteristic of this theory that, as far as we are aware, is reported here for the first time.

Finally, coming back to an even broader significance of the new solutions, one may wonder about possible consequences of the nonlocal terms if these are included in more general physical models that have a gauge structure (in Condensed Matter or High Energy Physics). It is also worth mentioning that if one follows the same "unconventional" method (of solution of PDEs) with the Maxwell's equations for the electric and magnetic fields (rather than with the PDEs for the potentials that give Λ), the corresponding nonlocal terms can be derived, and one can then see that these essentially demonstrate the causal propagation of

the radiation electric and magnetic fields outside physically inaccessible confined sources (i.e. solenoids or electric cages). Although this is of course widely known at the level of classical fields, a major conclusion that can be drawn from the present work (at the level of gauge transformations) is that *a corresponding Causality also exists at the level of quantum mechanical phases* as well, and this is enforced by the nonlocal terms in t-dependent cases. It strongly indicates that the nonlocal terms found in this work at the level of quantum mechanical phases reflect a causal propagation of wavefunction phases **in the Schrödinger Picture** (at least one part of them, the one containing the fields, which competes with the Aharonov-Bohm types of phases containing the potentials). This is an entirely new concept (given the *local nature* but also the *nonrelativistic character* of the Schrödinger equation) and deserves to be further explored. It would indeed be worth investigating possible applications of the above results (of nonlocal phases of wavefunctions, solutions of the local Schrödinger equation) in t-dependent single- *vs* double-slit experiments recently discussed by the group of Aharonov[9] who use a completely different method (with modular variables in the Heisenberg picture). One should also note recent work[10], that rightly emphasizes that Physics cannot currently predict how we dynamically go from the single-slit diffraction pattern to the double-slit diffraction pattern (whether it is in a gradual and causal manner or not) and where a relevant experiment is proposed to decide on (measure) exactly this. Application of our nonlocal terms in such questions in analogous experiments (i.e. by introducing (finite) scalar potentials on slits in a t-dependent way) provides a completely new formulation for addressing causal issues of this type, and is currently under investigation. Furthermore, *SU(2)* generalizations would be an obviously interesting extension of the above *U(1)* theory, and such generalizations are rather formally direct and not difficult to make (an immediate physically interesting question being whether the new nonlocal terms might have a nontrivial impact on i.e. spin-$\frac{1}{2}$-states, since these terms would act asymmetrically on opposite spins). Finally, it is worth noting that, if E's were substituted by gravitational fields and B's by Coriolis force fields arising in non-inertial frames of reference, the above nonlocalities (and their apparent causal nature) could possibly have an interesting story to tell about quantum mechanical phase behavior in a Relativistic/Gravitational framework.

Graduate students Kyriakos Kyriakou and Georgios Konstantinou of the University of Cyprus and Areg Ghazaryan of Yerevan State University are acknowledged for having care-

fully reproduced all (extremely long) results (most of them not reported here). Georgios Konstantinou is also acknowledged for having drawn Figures 2, 3 and 4. Dr. Cleopatra Christoforou of the Department of Mathematics and Statistics of the University of Cyprus is acknowledged for a discussion concerning the mathematical method followed.

A. Appendix A

The time-dependent Schrödinger equation for a particle (of mass m and charge q) moving in the set of potentials $(\mathbf{A}_1(\mathbf{r},t), \phi_1(\mathbf{r},t))$ is

$$\frac{[-\mathbf{i}\hbar\nabla - \frac{q}{c}\mathbf{A}_1(\mathbf{r},t)]^2}{2m}\Psi_1(\mathbf{r},t) + q\phi_1(\mathbf{r},t)\Psi_1(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\Psi_1(\mathbf{r},t) \tag{52}$$

and the one for the same particle moving in the set of potentials $(\mathbf{A}_2(\mathbf{r},t), \phi_2(\mathbf{r},t))$ is

$$\frac{[-\mathbf{i}\hbar\nabla - \frac{q}{c}\mathbf{A}_2(\mathbf{r},t)]^2}{2m}\Psi_2(\mathbf{r},t) + q\phi_2(\mathbf{r},t)\Psi_2(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\Psi_2(\mathbf{r},t). \tag{53}$$

Below we recall the general proof that solutions of the two above equations are formally connected through

$$\Psi_2(\mathbf{r},t) = e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}\Psi_1(\mathbf{r},t),$$

which is eq.(1) of the text, with the function $\Lambda(\mathbf{r},t)$ satisfying the system of PDEs (2), namely

$$\nabla\Lambda(\mathbf{r},\mathbf{t}) = \mathbf{A}_2(\mathbf{r},t) - \mathbf{A}_1(\mathbf{r},t) \qquad \mathit{and} \qquad -\frac{1}{c}\frac{\partial\Lambda(\mathbf{r},t)}{\partial t} = \phi_2\left(\mathbf{r},t\right) - \phi_1(\mathbf{r},t).$$

Indeed, it is an obvious vector identity that

$$[-\mathbf{i}\hbar\nabla - \frac{q}{c}(\mathbf{A}_1(\mathbf{r},t) + \nabla\Lambda(\mathbf{r},\mathbf{t}))]e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}\Psi_1(\mathbf{r},t) = e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}[-\mathbf{i}\hbar\nabla - \frac{q}{c}\mathbf{A}_1(\mathbf{r},t)]\Psi_1(\mathbf{r},t), \tag{54}$$

which, if applied once more (but now on the new function $\hat{Y}\ (\mathbf{r},t) = [-\mathbf{i}\hbar\nabla - \frac{q}{c}\mathbf{A}_1(\mathbf{r},t)]\Psi_1(\mathbf{r},t)$ in place of the single $\Psi_1(\mathbf{r},t)$) gives the well-known generalization

$$[-\mathbf{i}\hbar\nabla-\frac{q}{c}(\mathbf{A}_1(\mathbf{r},t)+\nabla\Lambda(\mathbf{r},\mathbf{t}))]^2 e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}\Psi_1(\mathbf{r},t)=e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}[-\mathbf{i}\hbar\nabla-\frac{q}{c}\mathbf{A}_1(\mathbf{r},t)]^2\Psi_1(\mathbf{r},t) \quad (55)$$

(and repeated application would of course give a similar identity for any positive integer power). In addition, we trivially have

$$i\hbar\frac{\partial}{\partial t}[e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}\Psi_1(\mathbf{r},t)]=e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}i\hbar\frac{\partial}{\partial t}\Psi_1(\mathbf{r},t)-e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}\frac{q}{c}\frac{\partial\Lambda(\mathbf{r},t)}{\partial t}\Psi_1(\mathbf{r},t). \quad (56)$$

One trivially notes then that, indeed, if $\Psi_2(\mathbf{r},t)$ is substituted by $e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}\Psi_1(\mathbf{r},t)$, $\mathbf{A}_2(\mathbf{r},t)$ is substituted by $\mathbf{A}_1(\mathbf{r},t)+\nabla\Lambda(\mathbf{r},\mathbf{t})$ and $\phi_2\,(\mathbf{r},t)$ is substituted by $\phi_1\,(\mathbf{r},t)-\frac{1}{c}\frac{\partial\Lambda(\mathbf{r},t)}{\partial t}$, then the left-hand-side (lhs) of (53) becomes $e^{i\frac{q}{\hbar c}\Lambda(\mathbf{r},t)}*$ [lhs of (52)$-\frac{q}{c}\frac{\partial\Lambda(\mathbf{r},t)}{\partial t}\Psi_1(\mathbf{r},t)$], while the right-hand-side of (53) is the above (56), and this equality is obviously satisfied (after cancellation of the $\frac{\partial\Lambda}{\partial t}$-additive term, and then of the common global phase factor from both sides) if $\Psi_1(\mathbf{r},t)$ satisfies (52). ✓

B. Appendix B

We present here the full derivation of the solutions of the system of PDEs (8), which if applied to only one spatial variable is (21), namely

$$\frac{\partial\Lambda(x,t)}{\partial x}=A(x,t) \qquad and \qquad -\frac{1}{c}\frac{\partial\Lambda(x,t)}{\partial t}=\phi\,(x,t)$$

(with $A(x,t)=A_2(x,t)-A_1(x,t)$ and $\phi\,(x,t)=\phi_2\,(x,t)-\phi_1\,(x,t)$). The system is underdetermined in the sense that we only have knowledge of Λ at an initial point (x_0,t_0) and with no further boundary conditions (hence multiplicities of solutions are generally expected, and these are discussed separately below and mainly in the text). Let us first look for unique (single-valued) solutions (i.e. with Λ being a *function* on the (x,t)-plane, in the sense of Elementary Analysis) and let us integrate the *first* of (21) – without dropping terms that may at first sight appear redundant – to obtain

$$\Lambda(x,t)-\Lambda(x_0,t)=\int_{x_0}^{x}A(x',t)dx'+\tau(t). \quad (57)$$

By then substituting this to the second of (21) (and assuming that interchanges of derivatives and integrals are allowed, i.e. covering cases of potentials with discontinuous first

derivatives, something that corresponds to the physical case of discontinuous magnetic fields – a case very often discussed in the literature), we obtain

$$\phi\left(x,t\right)=-\frac{1}{c}\int\limits_{x_0}^{x}\frac{\partial A(x',t)}{\partial t}dx'-\frac{1}{c}\frac{\partial\tau(t)}{\partial t}-\frac{1}{c}\frac{\partial\Lambda(x_0,t)}{\partial t}, \tag{58}$$

which if integrated gives

$$\tau(t)=\tau(t_0)+\Lambda(x_0,t_0)-\Lambda(x_0,t)-\int\limits_{t_0}^{t}dt'\int\limits_{x_0}^{x}dx'\frac{\partial A(x',t')}{\partial t'}-c\int_{t_0}^{t}\phi\left(x,t'\right)dt'+g(x) \tag{59}$$

with $g(x)$ to be chosen in such a way that the entire right-hand-side of (59) is only a function of t, as it should be (hence independent of x). Finally, by substituting $\frac{\partial A(x',t')}{\partial t'}$ with $-c\left(E(x',t')+\frac{\partial\phi(x',t')}{\partial x'}\right)$, (where $E(x',t')=E_2(x',t')-E_1(x',t')$), carrying out the integration with respect to x', and by demanding that $\tau(t)$ be independent of x, we finally obtain the following general solution

$$\Lambda(x,t)=\Lambda(x_0,t_0)+\int\limits_{x_0}^{x}A(x',t)dx'-c\int\limits_{t_0}^{t}\phi(x_0,t')dt'+\left\{c\int\limits_{t_0}^{t}dt'\int\limits_{x_0}^{x}dx'E(x',t')+g(x)\right\}+\tau(t_0)$$

with $g(x)$ chosen in such a way that the quantity $\left\{c\int\limits_{t_0}^{t}dt'\int\limits_{x_0}^{x}dx'E(x',t')+g(x)\right\}$ is independent of x. This is solution (22) of the text.

Here it should be noted that, if we had first integrated the *second* of (21) we would have

$$\Lambda(x,t)-\Lambda(x,t_0)=-c\int_{t_0}^{t}\phi(x,t')dt'+\chi(x) \tag{60}$$

and then from the first of (21) we would get

$$A\left(x,t\right)=-c\int\limits_{t_0}^{t}\frac{\partial\phi(x,t')}{\partial x}dt'+\frac{\partial\chi(x)}{\partial x}+\frac{\partial\Lambda(x,t_0)}{\partial x}, \tag{61}$$

which after integration would give

$$\chi(x)=\chi(x_0)+\Lambda(x_0,t_0)-\Lambda(x,t_0)+c\int\limits_{x_0}^{x}dx'\int\limits_{t_0}^{t}dt'\frac{\partial\phi(x',t')}{\partial x'}+\int_{x_0}^{x}A\left(x',t\right)dx'+\hat{g}(t) \tag{62}$$

with $\hat{g}(t)$ to be chosen in such a way that the entire right-hand-side of (62) is only a function of x, as it should be (hence independent of t). Finally, by substituting $\frac{\partial\phi(x',t')}{\partial x'}$ with $-\left(E(x',t')+\frac{1}{c}\frac{\partial A(x',t')}{\partial t'}\right)$, carrying out the integration with respect to t', and by demanding that $\chi(x)$ be independent of t, we would finally obtain the following general solution

$$\Lambda(x,t)=\Lambda(x_0,t_0)+\int_{x_0}^{x}A(x',t_0)dx'-c\int_{t_0}^{t}\phi\left(x,t'\right)dt'+\left\{-c\int_{x_0}^{x}dx'\int_{t_0}^{t}dt'E(x',t')+\hat{g}(t)\right\}+\chi(x_0)$$

with $\hat{g}(t)$ chosen in such a way that the quantity $\left\{-c\int_{x_0}^{x}dx'\int_{t_0}^{t}dt'E(x',t')+\hat{g}(t)\right\}$ is independent of t. This is solution (23) of the text.

Solutions (22) and (23) can be viewed as the (formal) analogs of (13) and (15) correspondingly, although they hide in them much richer Physics because of their dynamic character (see Section IX). (The additional constant last terms were shown in Section IX to be related to possible multiplicities of Λ, and they are zero in simple-connected spacetimes).

The reader was provided with the direct verification (i.e. proof by "going backwards") that (22) or (23) are indeed solutions of the basic system of PDEs (21) in Section VIII.

C. Appendix C

We provide here a general proof of the generalized Werner & Brill cancellations in simple-connected spacetime, namely, that solutions (22) and (23) are equivalent. By looking first at the general structure of solutions (22) and (23), we note that in both forms, the last constant terms ($\tau(t_0)$ and $\chi(x_0)$) are only present in cases where Λ is expected to be multivalued (this comes from the definitions of $\tau(t_0)$ and $\chi(x_0)$, see discussion below) and therefore these constant quantities are nonvanishing in cases of motion only in multiple-connected spacetimes (leading to phenomena of the Aharonov-Bohm type (see the analogous discussion in Appendix F, on the easier-to-follow magnetic case)). In such multiple-connected cases these last terms are simply equal (in absolute value) to the enclosed fluxes in regions of spacetime that are physically inaccessible to the particle (in the electric Aharonov-Bohm setup, for example, it turns out that $\tau(t_0)=-\chi(x_0)=$ enclosed "electric flux" in spacetime,

see below for the proof). Although such cases can also be covered by our method below, let us for the moment ignore them (set them to zero) and focus on cases of motion in simple-connected spacetimes. Then the two solutions (22) and (23) are actually equal as is now shown: since $\left\{c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E(x',t')+g(x)\right\}$ is independent of x, its x-derivative is zero which leads to $g'(x) = -c\int_{t_0}^{t}dt'E(x,t')$, with a general solution $g(x) = g(x_0) - c\int_{x_0}^{x}dx'\int_{t_0}^{t}dt'E(x',t')+C(t)$, and with a $C(t)$ such that the right-hand-side is only a function of x, hence independent of t; but this is exactly the form of (23), if we identify $C(t)$ with $\hat{g}(t)$ (and $g(x_0)$ with $\chi(x_0)$). This can be easily seen if we note that substitution of $E(x',t')$ with $-\frac{\partial\phi(x',t')}{\partial x'} - \frac{1}{c}\frac{\partial A(x',t')}{\partial t'}$ and two integrations carried out finally interchange the forms of the 1st solution (22) from $\left(\int_{x_0}^{x}A(x',t)dx' - c\int_{t_0}^{t}\phi(x_0,t')dt'\right)$ to $\left(\int_{x_0}^{x}A(x',t_0)dx' - c\int_{t_0}^{t}\phi(x,t')dt'\right)$ of the 2nd solution (23).

The above could alternatively be proven if in (59), instead of substituting $\frac{\partial A(x',t')}{\partial t'}$ in terms of the electric field difference, we had merely interchanged the ordering of integrations in the 1st integral term. This would then immediately take us to the 2nd solution (23), with automatically identifying the t-independent (hence x-dependent) quantity $\left\{-c\int_{x_0}^{x}dx'\int_{t_0}^{t}dt'E(x',t')+\hat{g}(t)\right\}$ of the 2nd solution (23) with the function $g(x)$ of the 1st solution (22). (In a similar way, one can prove the identification of the x-independent (hence t-dependent) quantity $\left\{c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E(x',t')+g(x)\right\}$ of the 1st solution (22) with the function $\hat{g}(t)$ of the 2nd solution (23)). This is the deep mathematical cause of the generalized Werner & Brill cancellations of the text in this electric case.

Finally, with respect to $\tau(t_0)$ and $\chi(x_0)$, let us give an example to see why ordinarily (in simple-connectivity) they are zero, or in the most general case they are related to physically inaccessible enclosed fluxes. Starting from (57), where $\tau(t)$ was first introduced, we have that

$$\tau(t_0) = \Lambda(x, t_0) - \Lambda(x_0, t_0) - \int_{x_0}^{x} A(x', t_0)dx', \tag{63}$$

which should be independent of x (and *it is* as can easily be proven, since its x-derivative gives $\frac{\partial\Lambda(x,t_0)}{\partial x} - A(x, t_0)$ which is zero, as $\Lambda(x, t)$ satisfies by assumption the first equation of the system (21) of PDEs (evaluated at $t = t_0$)). We can therefore determine its value by taking the limit $x \to x_0$ in (63), which is zero, unless there is a multivaluedness of Λ at the point (x_0, t_0). This happens for example for A having a δ-function form (a case however which we leave out, otherwise the assumed interchanges might not be allowed) or in cases that there is a "memory" that the system has multiplicities in Λ, i.e. in Aharonov-Bohm configurations (with enclosed and inaccessible fluxes in space-time), hence the value of $\tau(t_0)$ being expected to be equal to the enclosed "electric flux": the limit $x \to x_0$ (for fixed t_0) in the path sense of solution (22) that is needed then to determine $\tau(t_0)$ is equivalent to making an entire closed trip around the observation rectangle in the *positive* sense, landing on the initial point (x_0, t_0). This, from (63), gives that $\tau(t_0)$ = enclosed "electric flux". A similar argument applied for

$$\chi(x_0) = \Lambda(x_0, t) - \Lambda(x_0, t_0) + c\int_{t_0}^{t} \phi(x_0, t')dt' \tag{64}$$

leads to the value of $\chi(x_0)$ being equal to *minus* the enclosed "electric flux" (a corresponding limit $t \to t_0$ (for fixed x_0) in the path sense of solution (23) is now equivalent to making an entire trip around the rectangle in the *negative* sense, landing again on the same initial point (x_0, t_0)). This, from (64), gives that $\chi(x_0) = -$enclosed "electric flux". If these values are actually substituted in (22) (with $g(x) = 0$) and in (23) (with $\hat{g}(t) = 0$) they give the correct electric Aharonov-Bohm result (where effectively there are no nonlocal contributions, and only the line-integrals of A and ϕ contribute to the phase). [The above choice $g(x) = \hat{g}(t) = 0$ is a natural one, made because, in this Aharonov-Bohm case, the enclosed "electric flux" is independent of both x and t]. A more detailed discussion of such multiplicities in the standard *static* magnetic Aharonov-Bohm case is given in Appendix F.

D. Appendix D

After having discussed fully the simple (x,t)-case, let us for completeness give the analogous (Euclidian-rotated in 4-D spacetime) derivation for (x,y)-variables and briefly discuss the properties of the simpler static solutions, but now in full generality (also including possible multi-valuedness of Λ in the usual magnetic Aharonov-Bohm cases). We will simply need to apply the same methodology (of solution of a system of PDEs) to the system already shown in (14), namely

$$\frac{\partial\Lambda(x,y)}{\partial x} = A_x(x,y) \qquad \textit{and} \qquad \frac{\partial\Lambda(x,y)}{\partial y} = A_y(x,y).$$

By first integrating the 1st of this (again without dropping any terms that may appear redundant) we obtain the analog of (57), namely

$$\Lambda(x,y) - \Lambda(x_0,y) = \int_{x_0}^{x} A_x(x',y)dx' + f(y) \tag{65}$$

and by then substituting the result to the 2nd we have

$$A_y\,(x,y) = \int\limits_{x_0}^{x} \frac{\partial A_x(x',y)}{\partial y} dx' + f'(y) + \frac{\partial\Lambda(x_0,y)}{\partial y} \tag{66}$$

which if integrated leads to

$$f(y) = f(y_0) - \Lambda(x_0,y) + \Lambda(x_0,y_0) - \int\limits_{y_0}^{y} dy' \int\limits_{x_0}^{x} dx' \frac{\partial A_x(x',y')}{\partial y'} + \int_{y_0}^{y} A_y\,(x,y')\,dy' + g(x) \tag{67}$$

with $g(x)$ to be chosen in such a way that the entire right-hand-side of (67) is only a function of y, as it should be (hence independent of x). Finally, by substituting $\frac{\partial A_x(x',y')}{\partial y'}$ with $\frac{\partial A_y(x',y')}{\partial x'} - B_z(x',y')$, carrying out the integration with respect to x', and by demanding that $f(y)$ be independent of x, we finally obtain the following general solution

$$\Lambda(x,y) = \Lambda(x_0,y_0) + \int\limits_{x_0}^{x} A_x(x',y)dx' + \int\limits_{y_0}^{y} A_y(x_0,y')dy' + \left\{ \int\limits_{y_0}^{y} dy' \int\limits_{x_0}^{x} dx' B_z(x',y') + g(x) \right\} + f(y_0)$$

$$\textit{with} \;\; g(x) \;\; \textit{chosen so that} \;\; \left\{ \int\limits_{y_0}^{y} dy' \int\limits_{x_0}^{x} dx' B_z(x',y') + g(x) \right\} \textbf{:} \; \textsf{is independent of} \;\; x,$$

which is eq.(24) of the text, or eq.(13) but with included multiplicities through the extra $f(y_0)$ (which for simple-connected space can be set to zero). The result (24) applies to cases where the particle passes through *different* magnetic fields (recall that $B_z = (\boldsymbol{B}_2 - \boldsymbol{B}_1)_z$) in spatial regions that are remote to the observation point (x, y). Alternatively, by following the reverse route (first integrating the 2nd equation of the basic system (14)) we would obtain

$$\Lambda(x,y) - \Lambda(x,y_0) = \int_{y_0}^{y} A_y(x,y')dy' + \hat{h}(x) \tag{68}$$

and by then substituting the result to the 1st we would have

$$A_x\,(x,y) = \int_{y_0}^{y} \frac{\partial A_y(x,y')}{\partial x} dy' + \hat{h}'(x) + \frac{\partial \Lambda(x,y_0)}{\partial x} \tag{69}$$

which if integrated would lead to

$$\hat{h}(x) = \hat{h}(x_0) - \Lambda(x,y_0) + \Lambda(x_0,y_0) - \int_{x_0}^{x} dx' \int_{y_0}^{y} dy' \frac{\partial A_y(x',y')}{\partial x'} + \int_{x_0}^{x} A_x\,(x',y)\,dx' + h(y) \tag{70}$$

with $h(y)$ to be chosen in such a way that the entire right-hand-side of (70) is only a function of x, as it should be (hence independent of y). Finally, by substituting $\frac{\partial A_y(x',y')}{\partial x'}$ with $\frac{\partial A_x(x',y')}{\partial y'} + B_z(x',y')$, carrying out the integration with respect to y', and by demanding that $\hat{h}(x)$ be independent of y, we would finally obtain the following general solution

$$\Lambda(x,y) = \Lambda(x_0,y_0) + \int_{x_0}^{x} A_x(x',y_0)dx' + \int_{y_0}^{y} A_y(x,y')dy' + \left\{ -\int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x',y') + h(y) \right\} + \hat{h}(x_0)$$

$$\textit{with}\ \ h(y)\ \ \textit{chosen so that}\ \ \left\{ -\int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x',y') + h(y) \right\} : \text{is independent of}\ \ y,$$

which is eq.(25) of the text, or eq.(15) but with included multiplicities through the extra $\hat{h}(x_0)$. One can actually show that the two solutions are equivalent (i.e. (13) and (15) for a simple-connected space are equal[17]), a fact that can be proven in a way similar to the (x,t)-cases of Appendix C. (For the case of multiple-connectivity of the two-dimensional space, a brief discussion of the actual values of the multiplicities $f(y_0)$ and $\hat{h}(x_0)$ has been given at the end of Section X and is presented in more detail in Appendix F).

E. Appendix E

We here provide the spatial solutions in polar coordinates. By following a similar procedure (of solving the system of PDEs resulting from (9)) in polar coordinates (ρ, φ), namely

$$\frac{\partial \Lambda(\rho,\varphi)}{\partial \rho} = A_\rho(\rho,\varphi) \qquad and \qquad \frac{1}{\rho}\frac{\partial \Lambda(\rho,\varphi)}{\partial \varphi} = A_\varphi(\rho,\varphi)$$

with steps completely analogous to those of Appendix D, one can obtain the following analogs of solutions (24) and (25), namely

$$\Lambda(\rho,\varphi) = \Lambda(\rho_0,\varphi_0) + \int_{\rho_0}^{\rho} A_\rho(\rho',\varphi)d\rho' + \int_{\varphi_0}^{\varphi} \rho_0 A_\varphi(\rho_0,\varphi')d\varphi' + \left\{ \int_{\varphi_0}^{\varphi} d\varphi' \int_{\rho_0}^{\rho} \rho' d\rho' B_z(\rho',\varphi') + g(\rho) \right\} + f(\varphi_0) \tag{71}$$

$$with \;\; g(\rho) \;\; chosen \;\; so \;\; that \;\; \left\{ \int_{\varphi_0}^{\varphi} d\varphi' \int_{\rho_0}^{\rho} \rho' d\rho' B_z(\rho',\varphi') + g(\rho) \right\} : \text{ is independent of } \rho, \tag{72}$$

and

$$\Lambda(\rho,\varphi) = \Lambda(\rho_0,\varphi_0) + \int_{\rho_0}^{\rho} A_\rho(\rho',\varphi_0)d\rho' + \int_{\varphi_0}^{\varphi} \rho A_\varphi(\rho,\varphi')d\varphi' + \left\{ -\int_{\rho_0}^{\rho} \rho' d\rho' \int_{\varphi_0}^{\varphi} d\varphi' B_z(\rho',\varphi') + h(\varphi) \right\} + \hat{h}(\rho_0) \tag{73}$$

$$with \;\; h(\varphi) \;\; chosen \;\; so \;\; that \;\; \left\{ -\int_{\rho_0}^{\rho} \rho' d\rho' \int_{\varphi_0}^{\varphi} d\varphi' \in (\rho',\varphi') + h(\varphi) \right\} : \text{ is independent of } \varphi, \tag{74}$$

and in these, the proper choices of $g(\rho)$ and $h(\varphi)$ will again be determined by their corresponding conditions, depending on the actual shape of the B_z-distribution and the positioning of initial and final points (ρ_0, φ_0) and (ρ, φ). [Furthermore, the observation rectangle has now given its place to a slice of an annular section].

F. Appendix F

We here discuss the issue of multiplicities (the last terms of (24) and (25)) in case of spatial multiple-connectivity (such as the standard magnetic Aharonov-Bohm case, in which we can take $g(x) = 0$ *and* $h(y) = 0$, since the enclosed magnetic flux is independent of both x and y).

For the Aharonov-Bohm setting we will have to deal with multiple-connected space and with a (static) magnetic flux Φ being contained only in the physically inaccessible region. In such a case we know that the $\Lambda(\mathbf{r})$ that solves (9) is not single-valued. How is this fact (and the standard result (10)) compatible with the new formulation? To answer this in full generality we will consider two separate cases that arise naturally (pertaining to the issue of what the dummy variables (x', y') inside the B_z-terms of our results (i.e. of (13) and (15)) actually represent). First, if the variables x and y everywhere in the text always denote only coordinates of the region that is physically accessible to the particle, then B_z is everywhere vanishing, this effectively reducing (13) and (15) to

$$\Lambda(x, y) = \Lambda(x_0, y_0) + \int_{x_0}^{x} A_x(x', y)dx' + \int_{y_0}^{y} A_y(\mathbf{x_0}, y')dy' + C$$

$$\Lambda(x, y) = \Lambda(x_0, y_0) + \int_{x_0}^{x} A_x(x', \mathbf{y_0})dx' + \int_{y_0}^{y} A_y(x, y')dy' + C$$

with C a common constant; these are the standard results (the Dirac phases) along the two alternative paths discussed in the text (the red and green paths of Fig.1) that (through their difference) lead to the magnetic Aharonov-Bohm effect (Λ being no longer single-valued and the difference of the two solutions giving the enclosed (and physically inaccessible) Φ). Let us however be even more general and let us decide to use the variables x and y to *also* denote coordinates of the physically inaccessible region**;** this would be the case, if, for example, we had previously started with that region being accessible (i.e. through a penetrable scalar potential) and at the end we followed a limiting procedure (i.e. of this scalar potential going to infinity) so that this region would become in the limit impenetrable and therefore inaccessible. In such a case the variables x and y would now contain *remnants* of the previously allowed values (but currently not allowed for the description of particle coordinates) such as the values of the dummy variables x' and y' in the B_z-terms of (13) and (15)**;** such values would therefore still be present in the expressions giving Λ (even though

these dummy variables x' and y' would now describe an inaccessible region). In other words, the inaccessible B_z is still formally present in the problem and it shows up explicitly in the generalized gauge functions of the new formulation. How does this formulation then lead to the standard Aharonov-Bohm result in such a limiting case (essentially a case of smoothly-induced spatial multiple-connectivity)?

Before we answer this, the reader should probably be reminded that our formulation only deals with wavefunction-phases; questions therefore of rigid (vanishing) boundary conditions (on the boundary of the inaccessible region) that apply to (and must be imposed on) the entire wavefunction, and mostly on its modulus, can only be addressed indirectly (and as we will see, through a "memory" that the phases have of their multivaluedness, whenever the space is multiple-connected). To see this, we need the generalized results (eqs (24) and (25)) that contain the additional "multiplicities". These most general results (for multiple-connected space) were derived in Section IX and have the form

$$\Lambda(x,y) = \Lambda(x_0,y_0)+\int_{x_0}^{x} A_x(x',y)dx'+\int_{y_0}^{y} A_y(x_0,y')dy'+\left\{\int_{y_0}^{y} dy'\int_{x_0}^{x} dx' B_z(x',y') + g(x)\right\}+f(y_0)$$

and

$$\Lambda(x,y) = \Lambda(x_0,y_0)+\int_{x_0}^{x} A_x(x',y_0)dx'+\int_{y_0}^{y} A_y(x,y')dy'+\left\{-\int_{x_0}^{x} dx'\int_{y_0}^{y} dy' B_z(x',y') + h(y)\right\}+\hat{h}(x_0)$$

with the functions $g(x)$ and $h(y)$ satisfying the same conditions as in (13) and (15). We note the extra appearance of the new constant terms $f(y_0)$ and $\hat{h}(x_0)$ (the "multiplicities") and these are "defined" (see (65) and (68) where the functions f and $\hat{h}$ were first introduced) by

$$f(y_0) = \Lambda(x,y_0) - \Lambda(x_0,y_0) - \int_{x_0}^{x} A_x(x',y_0)dx'$$

and

$$\hat{h}(x_0) = \Lambda(x_0,y) - \Lambda(x_0,y_0) - \int_{y_0}^{y} A_y(x_0,y')dy'.$$

Let us then identify proper choices for the functions $g(x)$ and $h(y)$ and for the constants $f(y_0)$ and $\hat{h}(x_0)$ in the above case of spatial multiple-connectivity (such as the standard magnetic Aharonov-Bohm case, with a non-extended (and static) magnetic flux in the forbidden region): First, we can always take $g(x) = 0$ and $h(y) = 0$ (always up to a common

constant as discussed earlier), since the enclosed magnetic flux is (in this Aharonov-Bohm case) independent of both x and y – the conditions of $g(x)$ and $h(y)$ being then automatically satisfied. Second, let us look more closely at the above "definitions" of $f(y_0)$ and $\hat{h}(x_0)$ **:**

We first note that $f(y_0)$ must be independent of x, and this is indeed true as is apparent by formally taking the derivative of the above definition of $f(y_0)$ with respect to x**;** we then have $\frac{\partial f(y_0)}{\partial x} = \frac{\partial \Lambda(x,y_0)}{\partial x} - A_x(x,y_0)$ which is indeed zero (as $\Lambda(x,y)$ satisfies by assumption the first equation of the system (14) of PDEs (evaluated at $y = y_0$)), showing that $\frac{\partial f(y_0)}{\partial x} = 0$ and that $f(y_0)$ does not really depend on the variable x that appears in its definition. We can therefore determine its value by taking the limit $x \to x_0$ (for fixed y_0)**:** we see from the above that this limit is simply equal to $\lim_{x\to x_0} \Lambda(x,y_0) - \Lambda(x_0,y_0)$ [as mentioned earlier, we leave out cases where A_x has a δ-function form, so that interchanges of all integrals are allowed in our earlier derivations (in Appendix D)], and this difference is nonzero only when there is a multivaluedness of Λ at the point (x_0,y_0), as *is* actually our case. The limit $x \to x_0$ (for fixed y_0) in the path-sense of solution (13) (or of (24)) that is then needed here in order to determine $f(y_0)$, is equivalent to making an entire closed trip around the observation rectangle in the *negative* sense, landing on the initial point (x_0,y_0), this therefore giving the value $f(y_0) =$ minus enclosed magnetic flux $= -\Phi$ (which is indeed a constant independent of x and y, as it *should* be). By following a completely symmetric argument for the above definition of $\hat{h}(x_0)$ (and by now taking the limit $y \to y_0$ (for fixed x_0), that is now equivalent to going around the loop in the *positive* sense, landing again on the initial point (x_0,y_0)) we obtain $\hat{h}(x_0) = +\Phi$. If these values of $f(y_0)$ and $\hat{h}(x_0)$ are finally substituted in the above most general solutions (eqs (24) and (25)) together with $g(x) = h(y) = 0$, then we note that $f(y_0)$ cancels out the $\int\limits_{y_0}^{y} dy' \int\limits_{x_0}^{x} dx' B_z(x',y')$ term (which is here just equal to the inaccessible flux Φ), and $\hat{h}(x_0)$ cancels out the $-\int\limits_{x_0}^{x} dx' \int\limits_{y_0}^{y} dy' B_z(x',y')$ term, and the two solutions are then once again reduced to the usual solutions of mere A-integrals along the two paths (i.e. the standard Dirac phase, with no nonlocal contributions) – their difference giving the closed loop integral of $\boldsymbol{A}$, hence the inaccessible flux and, finally, the well-known magnetic Aharonov-Bohm result. One should note here that the standard result in the new formulation requires some effort and it is only derived indirectly (due to the fact that we only deal with phases and not the moduli of wavefunctions, on which boundary conditions

are normally imposed), and it basically comes from the "memory" of the multivaluedness that the "gauge function" Λ carries (due to the multiple-connectivity of space).

G. Appendix G

We here present the method of solution of the system (26), and provide a detailed derivation of solutions (28), (32), (33) and (37) of the text. Starting with the *second* of (26), and by integrating it we obtain the expected generalization of (68), namely

$$\Lambda(x,y,t)-\Lambda(x,y_0,t)=\int_{y_0}^{y}A_y(x,y',t)dy'+f(x,t) \tag{75}$$

which if substituted to the first of (26) gives (after integration over x') a t-generalization of (70), namely

$$f(x,t)=f(x_0,t)-\Lambda(x,y_0,t)+\Lambda(x_0,y_0,t)-\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'\frac{\partial A_y(x',y',t)}{\partial x'}+\int_{x_0}^{x}A_x\left(x',y,t\right)dx'+G(y,t) \tag{76}$$

with $G(y,t)$ to be chosen in such a way that the entire right-hand-side of (76) is only a function of x and t, as it should be (hence independent of y). Finally, by substituting $\frac{\partial A_y(x',y',t)}{\partial x'}$ with $\frac{\partial A_x(x',y',t)}{\partial y'}+B_z(x',y',t)$, carrying out the integration with respect to y', and by demanding that $f(x,t)$ be independent of y, we obtain the following temporal generalization of (25)

$$\Lambda(x,y,t)=\Lambda(x_0,y_0,t)+\int_{x_0}^{x}A_x(x',y_0,t)dx'+\int_{y_0}^{y}A_y(x,y',t)dy'+$$

$$+\left\{-\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y',t)+G(y,t)\right\}+f(x_0,t)$$

$$\textit{with}\ \ G(y,t)\ \ \textit{such that}\quad \left\{-\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y',t)+G(y,t)\right\}:\ \ \text{is independent of}\ \ y.$$

This is eq.(27) of the text. From this point on, the third equation of the system (26) is getting involved to determine the nontrivial effect of scalar potentials on $G(y,t)$; by combining it with (27) there results a wealth of patterns: integration with respect to t' leads to

$$G(y,t)=G(y,t_0)-\Lambda(x_0,y_0,t)+\Lambda(x_0,y_0,t_0)-f(x_0,t)+f(x_0,t_0)-c\int_{t_0}^{t}\phi(x,y,t')dt'-$$

$$-\left[\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'\frac{\partial A_x(x',y_0,t')}{\partial t'}+\int_{t_0}^{t}dt'\int_{y_0}^{y}dy'\frac{\partial A_y(x,y',t')}{\partial t'}\right]+\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'\frac{\partial B_z(x',y',t')}{\partial t'}+F(x,y) \tag{77}$$

with $F(x,y)$ to be chosed in such a way that the entire right-hand-side of (77) is only a function of (y,t), as it should be, hence independent of x. In (77) there are two possible ways to determine the term in brackets, and another two ways to determine the term containing B_z. The easiest to follow (the one that more directly leads to the final *conditions* that the functions $F(x,y)$ and $G(y,t_0)$ are required to satisfy) is**:** (i) to substitute $\frac{\partial A_x(x',y_0,t')}{\partial t'}$ with $-c\left(E_x(x',y_0,t')+\frac{\partial\phi(x',y_0,t')}{\partial x'}\right)$ (and similarly for $\frac{\partial A_y(x,y',t')}{\partial t'}$), and (ii) to use the proviso that magnetic and electric fields are connected through the Faraday's law of Induction, namely $\frac{\partial B_z(x',y',t')}{\partial t'}=-c\left(\frac{\partial E_y(x',y',t')}{\partial x'}-\frac{\partial E_x(x',y',t')}{\partial y'}\right)$. These substitutions lead to cancellations of several intermediate quantities in (27) and (77) and lead to the final result (which is eq.(28) of the text), namely

$$\Lambda(x,y,t)=\Lambda(x_0,y_0,t_0)+\int_{x_0}^{x}A_x(x',y_0,t)dx'+\int_{y_0}^{y}A_y(x,y',t)dy'-\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y',t)+G(y,t_0)-$$

$$-c\int_{t_0}^{t}\phi(x_0,y_0,t')dt'+c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E_x(x',y,t')+c\int_{t_0}^{t}dt'\int_{y_0}^{y}dy'E_y(x_0,y',t')+F(x,y)+f(x_0,t_0)$$

with the functions $G(y,t_0)$ and $F(x,y)$ to be chosen in such a way as to satisfy the following 3 independent conditions**:**

$$\left\{G(y,t_0)-\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y',t_0)\right\}\textbf{:}\ \textit{is}\ \textsf{independent of}\ \ y,$$

which is eq.(29) of the text (and a special case of the condition on $G(y,t)$ above (see after (27))), and the other 2 turn out to be of the form

$$\left\{ F(x,y) + c\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' E_x(x',y,t') \right\} : \textit{is} \text{ independent of } x,$$

$$\left\{ F(x,y) + c\int_{t_0}^{t} dt' \int_{y_0}^{y} dy' E_y(x,y',t') \right\} : \textit{is} \text{ independent of } y,$$

which are eqs (30) and (31) of the text. It should be noted (for the reader who wants to follow all the steps) that the final condition (31) does *not* come out *directly* as the other two; because the function $G(y,t)$ has disappeared from the final form (28), one needs to *separately* impose the condition above for $G(y,t)$ $\Bigg($namely $\left\{ -\int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x',y',t) + G(y,t) \right\}$: independent of $y\Bigg)$ directly on the form (77); and in so doing, it is advantageous to interchange integrations (namely, do the t'-integral first) in the B_z-term of (77), so that $\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' \int_{y_0}^{y} dy' \frac{\partial B_z(x',y',t')}{\partial t'} = \int_{x_0}^{x} dx' \int_{y_0}^{y} dy' (B_z(x',y',t) - B_z(x',y',t_0))$, and then impose the (less stringent) condition (29) on $G(y,t_0)$; by following this strategy, after a number of cancellations of intermediate quantities one finally obtains the 3rd condition (31) on $F(x,y)$. (As for the constant quantity $f(x_0,t_0)$ appearing in (28), this again describes possible effects of multiple-connectivity at the instant t_0 (which are absent for simple-connected spacetimes, but were crucial in the discussion of the van Kampen thought-experiment of the text)).

Eq. (28) is our first solution. It is now crucial to note that an alternative form of solution (with the functions $G's$ and F satisfying the *same* conditions as above) can be derived if, in the term in brackets of (77) we merely interchange integrations, leaving therefore A's everywhere rather than introducing electric fields; following at the same time the above strategy of changing the ordering of integrations in the B_z-term as well (without therefore using Faraday's law) this alternative form of solution turns out to be

$$\Lambda(x,y,t) = \Lambda(x_0,y_0,t_0) + \int_{x_0}^{x} A_x(x',y_0,t)dx' + \int_{y_0}^{y} A_y(x,y',t)dy' - \int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x',y',t_0) + G(y,t_0) -$$

$$-c\int_{t_0}^{t} \phi(x_0,y_0,t')dt' + c\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' E_x(x',y_0,t') + c\int_{t_0}^{t} dt' \int_{y_0}^{y} dy' E_y(x,y',t') + F(x,y) + f(x_0,t_0),$$

and is eq.(32) of the text. In this alternative solution we note that, in comparison with (28), the line-integrals of $\boldsymbol{E}$ have changed to the *other* alternative "path" (note the difference in the placement of the coordinates of the initial point (x_0, y_0) in the arguments of E_x and E_y) and they happen to have the same sense as the $\boldsymbol{A}$-integrals, while simultaneously the magnetic flux difference shows up with its value at the initial time t_0 rather than at t. This alternative form is useful in cases where we want to directly compare physical situations in the present (at time t) and in the past (at time t_0), and the above noted change of sense of $\boldsymbol{E}$-integrals (compared to (28)) was crucial in the discussion of the text.

Once again the reader can directly verify that (28) or (32) indeed satisfy the basic input system (26). (This verification is considerably more tedious than the ones of the main text but straightforward, and is not shown here).

But in order to make the above formalism useful for the van Kampen case, namely an enclosed (and physically inaccessible) magnetic flux (which however is *time-dependent*), it is important to have the analogous forms through a reverse route, namely starting with (integrating) the *first* of (26) and then substituting the result to the second**;** in this way we will at the end have the reverse "path" of $\boldsymbol{A}$-integrals, so that by taking the *difference* of the resulting solution and the above solution (28) (or (32)) will lead to the *closed* line integral of $\boldsymbol{A}$ which will be immediately related to the van Kampen's magnetic flux (at the instant t). By following then this route, and by applying a similar strategy at every intermediate step, we finally obtain the following solution (the spatially "dual" of (28)), which is eq.(33) of the text, namely

$$\Lambda(x,y,t) = \Lambda(x_0,y_0,t_0) + \int_{x_0}^{x} A_x(x',y,t)dx' + \int_{y_0}^{y} A_y(x_0,y',t)dy' + \int_{x_0}^{x} dx' \int_{y_0}^{y} dy' B_z(x',y',t) + \hat{G}(x,t_0) -$$

$$-c\int_{t_0}^{t} \phi(x_0,y_0,t')dt' + c\int_{t_0}^{t} dt' \int_{x_0}^{x} dx' E_x(x',y_0,t') + c\int_{t_0}^{t} dt' \int_{y_0}^{y} dy' E_y(x,y',t') + F(x,y) + \hat{h}(y_0,t_0)$$

with the functions $\hat{G}(x,t_0)$ and $F(x,y)$ to be chosen in such a way as to satisfy the following 3 independent conditions**:**

$$\left\{\hat{G}(x,t_0)+\int_{y_0}^{y}dy'\int_{x_0}^{x}dx'B_z(x',y',t_0)\right\}:\ is\ \text{independent of}\ \ x,$$

$$\left\{F(x,y)+c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E_x(x',y,t')\right\}:\ is\ \text{independent of}\ \ x,$$

$$\left\{F(x,y)+c\int_{t_0}^{t}dt'\int_{y_0}^{y}dy'E_y(x,y',t')\right\}:\ is\ \text{independent of}\ \ y,$$

where again for the above results the Faraday's law was crucial. The above conditions are eqs (34)-(36) of the text.

Finally, the corresponding analog of the alternative form (32) (i.e. with B_z now appearing at t_0) is more important and turns out to be

$$\Lambda(x,y,t)=\Lambda(x_0,y_0,t_0)+\int_{x_0}^{x}A_x(x',y,t)dx'+\int_{y_0}^{y}A_y(x_0,y',t)dy'+\int_{x_0}^{x}dx'\int_{y_0}^{y}dy'B_z(x',y',t_0)+\hat{G}(x,t_0)-$$

$$-c\int_{t_0}^{t}\phi(x_0,y_0,t')dt'+c\int_{t_0}^{t}dt'\int_{x_0}^{x}dx'E_x(x',y,t')+c\int_{t_0}^{t}dt'\int_{y_0}^{y}dy'E_y(x_0,y',t')+F(x,y)+\hat{h}(y_0,t_0)$$

which is eq.(37) of the text, with $\hat{G}(x,t_0)$ and $F(x,y)$ following the same 3 conditions above. The constant term $\hat{h}(y_0,t_0)$ again describes possible multiplicities at the instant t_0; it is absent for simple-connected spacetimes, but was crucial in the discussion of the van Kampen thought-experiment of Section XII.

In (33) (and in (37)), note the "alternative paths" (compared to solution (28) (and (32))) of line integrals of $\boldsymbol{A}$'s (or of $\boldsymbol{E}$'s). But the most crucial element for what is done in the text is the exclusive use of forms (32) and (37) (where B_z only appears at t_0), and the fact that, within each solution, the sense of $\boldsymbol{A}$-integrals is the *same* as the sense of the $\boldsymbol{E}$-integrals. (This is *not* true in the other solutions where $B_z(..,t)$ appears, as the reader can directly see). These facts were crucial to the discussion that addresses the so called van Kampen "paradox" in Section XII. It is the subtraction of these two forms (32) and (37) (where the B_z-term is evaluated at t_0) that leads to the final **causal** result that $\Delta\Lambda(t)=\Phi(t_0)$ in the

main text (when the spacetime point (x, y, t) is such that, at the instant t, the physical information (generated at t_0) has not yet reached the spatial point (x, y)).

---

[1] Y. Aharonov and D. Bohm, *Phys. Rev.* **115**, 485 (1959)

[2] S. Popescu, *Nature Physics* **6**, 151 (2010)

[3] An error often made (when the 1-D path is a closed loop, such as a nanoring) is the simultaneous imposition of spatial single-valuedness on both $\Psi(x,t)^{(\Phi)}$ and $\Psi(x,t)^{(0)}$ leading (erroneously) to quantization of the enclosed inaccessible flux Φ. Obviously Φ can have any arbitrary value, and spatial single-valuedness must only be imposed on $\Psi(x,t)^{(\Phi)}$

[4] H. R. Brown & P. R. Holland, *Amer. Journ. Phys.* **67**, 204 (1999)

[5] Most of the main results of this work have been presented in 2 Conferences in Tel Aviv and in Bristol ("*50 Years of the Aharonov-Bohm Effect and 25 Years of the Berry's Phase*") and, after having gone through close scrutiny, they have appeared in the special issue of *Journ. Phys. A: Math. Theor.* **43**, 354019 (2010)

[6] H. Weyl, *Z. Phys.* **56**, 330 (1929); for a historical review see L. O'Raifeartaigh & N. Straumann, *Rev. Mod. Phys.* **72**, 1 (2000)

[7] For a general presentation of multivalued gauge transformations see H. Kleinert, "*Multivalued Fields in Condensed Matter, Electromagnetism and Gravitation*", World Scientific (2008)

[8] M. Peshkin & A. Tonomura, Lecture Notes in Physics 340, Springer-Verlag (1989), Part I

[9] J. Tollaksen, Y. Aharonov, A. Casher, T. Kaufherr, and S. Nussimov, *New Journal of Physics* **12**, 013023 (2010)

[10] G. P. He, arXiv:0907.1974 (Nov. 2009)

[11] F. G. Werner & D. R. Brill, *Phys. Rev. Lett.* **4**, 344 (1960)

[12] N. G. van Kampen, *Phys. Lett.* **106A**, 5 (1984)

[13] T. T. Wu and C. N. Yang, *Phys. Rev. D* **12**, 3845 (1975)

[14] P. D. Noerdlinger, *Il Nuovo Cimento* **23**, 158 (1962)

[15] R. A. Brown & D. Home, *Il Nuovo Cimento* **107B**, 303 (1992)

[16] T. Troudet, *Phys. Lett.* **111A**, 274 (1985)

[17] This has been first noted by graduate student K. Kyriakou

[18] R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman Lectures on Physics*, Addison-Wesley (1964), vol. II, Chapter 15

[19] B. Felsager, "*Geometry, Particles, and Fields*", Springer-Verlag (1998), p.55

[20] H. Batelaan & A. Tonomura, *Physics Today* **62** (issue no. 9), p. 38 (September 2009)

[21] P. G. Luan & C. S. Tang, *J. Phys.: Condens. Matt.* **19**, 176224 (2007)

[22] *Compendium of Quantum Physics*, Ed. D. Greenberger *et al.*, Springer-Verlag 2009

[23] H. Erlichson, *Amer. Journ. Phys.* **38**, 162 (1970)

[24] I.e. M. P. Silverman, "*More than One Mystery*", Springer-Verlag (1995), or, better, the same author's more recent "*Quantum Superposition*", Springer-Verlag (2008), p.13

[25] Ref.[18], p. 15-13

[26] First book of Ref.[24], p. 16**;** or second book of Ref.[24], p. 19

FIGURE CAPTIONS

Figure 1. (Color online)**:** Examples of simple field-configurations (in simple-connected regions), where the nonlocal terms exist and are nontrivial, but can easily be determined**:** (a) a striped case in 1+1 spacetime, where the electric flux enclosed in the "observation rectangle" is dependent on t but independent of x; (b) a triangular distribution in 2-D space, where the part of the magnetic flux inside the corresponding "observation rectangle" depends on *both* x *and* y. The appropriate choices for the corresponding nonlocal functions $g(x)$ and $\hat{g}(t)$ for case (a), or $g(x)$ and $h(y)$ for case (b), are given in the text (Sections VIII and VI respectively).

Figure 2. (Color online)**:** The analog of paths of Fig.1 but now in 2+1 spacetime for the van Kampen thought-experiment, when the instant of observation t is so short that the physical information has not yet reached the spatial point of observation (x, y). The two solutions (that, for wavepackets, have to be subtracted in order to give the phase difference at (x, y, t)) are eqs (32) and (37) of the text, and are here characterized through their electric field E-line integral behavior**:** "electric field path (I)" (the red-arrow route) denotes solution (37), and "electric field path (II)" (the green-arrow route) denotes solution (32). Note that the strips of Fig.1(a) have now given their place to a light-cone. At the point of observation, the Aharonov-Bohm phase difference has now become "causal" due to cancellations between the two solutions (the two "electric field paths" above).

Figure 3. (Color online)**:** The standard double-slit apparatus with an additional strip

of a perpendicular magnetic field B of width W placed between the slit-region and the observation screen. In the text we deal for simplicity with the case $W << L$, so that deflections (of the semiclassical trajectories) due to the Lorentz force, shown here for a negative charge q, are very small.

Figure 4. (Color online)**:** The analog of Fig.3 (again for a negative q) but with an additional electric field parallel to the observation screen that is turned on for a time interval T. In the text we deal for simplicity with the case $T << \frac{L}{v}$ (with $v = \frac{1}{m}\frac{h}{\lambda}$, λ the de Broglie wavelength), so that deflections (of the semiclassical trajectories) due to the electric force are again very small. For both Fig.3 and 4, it is shown in the text that $\Delta\varphi^{semiclassical} = -\Delta\varphi^{AB}$, hence we observe an extra minus sign compared to what is usually reported in the literature.